REMARKS

Applicants request entry of the amendments and reconsideration of the application.

Amendment to the Specification

The amendment to the specification corrects typographical or inadvertent errors and does not add new matter. As noted in the Appendix A, which shows the changes made, the errors relate to the abbreviated titles of journal references and the form and completeness of the citations to these journal references.

Claims 17-36 have been canceled and new claims 37-76 have been added.

The amendments to the claims reflect the selection of SEQ ID NO: 17 and 18 for examination, as required by the Examiner in his objections to the claims (see page 4 of Paper No. 19). Applicants specifically reserve the right to seek patents for all the sequences and subject matter disclosed in the application and original claims. The changes in the recitation of the polynucleotides of the claims are made to better reflect the disclosures in the specification and in the priority document. The amendments are not made for reasons of patentability.

As shown below, no new matter enters by the new claims and the specification enables the new claims.

The new claims are directed to transcription factors that have at least 42% amino acid sequence identity with the AP2 transcription factor of SEQ ID NO: 18, and have enhanced plant disease tolerance due to changed expression or activity of the transcription factor. These claims are disclosed and enabled by the text of the specification at, for example, page 17, lines 6-9; page 13, lines 13-23; page 22, lines 34-35; page 23, lines 36-37 and continued on page 24, lines 1-4; page 25, lines 6-11; at page 11, lines 21-22, which discloses "[t]ranscription factors that are homologs of the disclosed sequences will typically share at least 40% amino acid sequence identity"; and at page 25, lines 4-9, which discloses that G19 (SEQ ID NO:22) expression was significantly induced upon infection by the fungal pathogen *Erysiphe* and G19-overexpressing plants were more tolerant to *Erysiphe* infection; and by Figure 1, which discloses the AP2-family transcription factors G1006 (SEQ ID NO: 34), G19 (SEQ ID NO:22), G4 (SEQ ID NO:12) and

G25 (SEQ ID NO: 34), which, as shown below, share 60%, 42%, 42%, and 54% amino acid sequence identity with G28 (SEQ ID NO:18), respectively.

As shown in the attached Appendix B providing experimental expression data, G1006 and G25 expression is induced in response to exposure to different disease agents. As shown in the executed declaration of Dr. Jacqueline Heard (the Heard Declaration, attached as Appendix C), G1006, G19, G4 and G25 expression has been shown to be up-regulated by *Fusarium, Erysiphe, Botrytis* exposure or infection, and methyl jasmonate in microarray analyses (see in particular the Heard Declaration, Exhibits A through G of the Declaration). As disclosed in the specification and stated in the Heard Declaration, the jasmonic acid signal transduction pathway is involved in the regulation of the defense response; the regulation of G4 expression by both hormones, its induction upon *Erysiphe orontii* infection, as well as the preliminary data indicating that increased tolerance to that pathogen is conferred by G19 overexpression, suggest that these transcription factors control the defense response. Thus, evidence is provided that shows AP2 family transcription factors with at least 42% identity to SEQ ID NO:18 (G28) either induce enhanced tolerance to plant disease or provide enhanced tolerance based on sequence similarity and expression profiles, thus likely endowing a plant with increased pathogen resistance (as would be recognized by one skilled in the art; see the Heard declaration, sections 8 to 9).

As also shown in Heard Declaration, the function of SEQ ID NOs:17 and 18 (G28) was more recently analyzed using transgenic plants in which this gene was expressed under the control of the 35S promoter. G28 overexpressing lines were shown to be more tolerant to infections by *Erysiphe orontii*, *Sclerotinia sclerotiorum* or *Botrytis cinerea*. Thus, the disclosure in the specification, that the transcription factors recited in the claims, including G28, could be used to produce plants that are more tolerant to disease, has been confirmed by experimentation. (See, for example, page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13.)

The new claims are also directed to transcription factors that comprise a conserved domain of a plant AP2 transcription factor wherein the conserved domain comprises one of an amino acid sequence of residues 145-213 of the AP2 transcription factor of SEQ ID NO: 18, or

an amino acid sequence having at least 84% identity to residues 145-213 of the AP2 transcription factor of SEQ ID NO: 18, or an amino acid sequence of residues 145-213 of the AP2 transcription factor of SEQ ID NO: 18 having one or more conservative substitutions, deletions, or insertions. Applicants clearly identified amino acid residues 145-213 of the AP2 transcription factor of SEQ ID NO: 18 in Figure 1a (see SEQ ID NO: 18, G28, in particular). Also, the specification at page 10, lines 13-14, specifically notes the conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for residues that may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority. In addition, fragments of domains of transcription factor sequences are specifically disclosed at page 4, lines 32-37, and in Figure 1.

Applicants also note that SEQ ID NO: 22 (G19) is disclosed in the specification as being significantly induced upon infection by the fungal pathogen *Erysiphe orontii* (*see* specification at page 25, lines 4-9). As shown in the accompanying BLAST report (Appendix F), the conserved domains of G28 (SEQ ID NO: 18) show substantial sequence identity to G19 (SEQ ID NO: 22), namely 60% identity (see Appendix F: BLAST report, page number listed as 12 or 28; handwritten notes regarding % identity provided by the applicants). Applicants have inserted this identity language into the new claims. The use of identity ranges is well known in the art and is specifically supported by the specification at, for example, page 11, lines 21-23. Thus, one of skill in the art would recognize the ability of the G19 sequence (SEQ ID NO: 22) and the closely related G28 sequence (SEQ ID NO: 18) to produce transgenic plants with at least pathogen stress resistance or disease resistance.

The BLAST report also shows that conserved domains of G4 (SEQ ID NO: 12) with 62% identity to SEQ ID NO: 18, and G1006 (SEQ ID NO: 34) with 97% identity to SEQ ID NO: 18, and G25 (SEQ ID NO: 82) with 69% identity, show substantial sequence identity to the conserved domain of G28 (SEQ ID NO: 18) (*see* handwritten notes regarding % identity provided by the applicants, particularly at pages 4, 10, and 13, respectively). The BLAST analysis method is fully disclosed in the specification at page 11, lines 3-20; at page 19, lines 2-8; and at page 25, lines 32-37. In addition, the specification discloses the function and utility of

D

these polynucleotide and polypeptide sequences (see specification at pages 1 through 3, and pages 5 through 8, where the different polynucleotide and polypeptide sequences and their use are disclosed). Accordingly, applicants' specification, in light of this evidence, clearly demonstrates the range of polynucleotide sequences that could be used in the new methods and transgenic plants.

To further demonstrate that the functional characteristics recited in the new claims would have been understood to be present with the recited polynucleotides, and/or easily identified in transgenic plants comprising these polynucleotides, applicants also enclose a copy of Table 1 of U.S. provisional application serial no. 60/125,814, at page 46 (see Appendix D, p 45, lines 15-32; page 46, Table 1). The present application claims priority to this provisional application. Table 1 shows amino acid residues of a polypeptide sequence that may be substituted with other amino acids and are restricted to conservative substitutions.

The conserved domains of AP2-family transcription factors are more nearly identical in their sequences, with G1006 (SEQ ID NO: 34), G19 (SEQ ID NO: 22), G4 (SEQ ID NO: 12) and G25 (SEQ ID NO: 82) sharing 98%, 72%, 69%, and 72% sequence identity, respectively, with the conserved domain of G28 (please see Appendix E, showing each of these comparisons and the sequence identity). Thus, structurally and functionally related AP2 family transcription factors share a minimum of 69% sequence identity in their conserved domains, which supports new claims directed to conserved domains.

Support for claim 37 is provided by claim 1 as filed (transgenic plant comprising recombinant polynucleotide altering disease resistance or tolerance), and the specification at page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests), and by page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity).

Support for claim 38 may be found in Figure 1a (conserved domain comprises residues 145-213), and the specification at page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity), and on page 6, lines 4-6 ("conserved domain" refers

D

to a polypeptide fragment that is more conserved at a sequence level than other fragments). The specification at page 10, lines 12-26 and page 10, lines 13-14 also notes conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for residues that may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority.

Support for new claim 39 is provided by the specification at, for example, page 5, line 9, and page 16, line 9 through page 17, line 24.

Support for new claim 40 is provided by claim 3 as filed and on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 41 is provided by claim 3 as filed and on page 13, lines 15-19.

Support for new claim 42 is provided by claim 4 as filed and on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 43 is provided on, for example, page 17, lines 10-11 (utility for increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, line 32 through page 21, line 28 (those genes induced after exposure to biotropic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claim 44 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

Support for new claim 45 is provided by claim 5 as filed (method for altering the disease tolerance or resistance of a plant), and, for example, on page 7, line 35 (modulating a plant's response to disease), page 17, lines 10-23 (increasing tolerance or resistance to pathogens and pests), and on page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity).

Support for new claim 46 is provided in Figure 1a (conserved domain comprises residues 145-213), on page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity), on page 6, lines 4-6 ("conserved domain" refers to a polypeptide fragment that is more conserved at a sequence level than other fragments), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests). The specification at page 10, lines 12-26 and page 10, lines 13-14 also notes conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for residues that may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority.

Support for new claim 47 is provided by, for example, page 5, line 9, page 6, lines 28-31, and page 17, lines 10-23.

Support for new claim 48 is provided by claim 8 as filed and on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 49 is provided by claim 8 as filed and on page 13, lines 15-19.

Support for new claim 50 is provided by claim 9 as filed and on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 51 is provided on, for example, page 17, lines 10-11 (increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, lines 32 through page 21, line 28 (those genes induced after exposure to biotropic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claims 52 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

D

Support for new claim 53 is provided by claim 10 as filed (method for altering the expression levels of at least one gene in a plant), and for example, on page 2, lines 19-23 and page 8, lines 29-34 (producing transgenic plants with modified expression levels of at least one gene), on page 17, lines 10-23 (increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests), and on page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity).

Support for new claim 54 may be found in Figure 1a (conserved domain comprises residues 145-213), on page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity), and on page 6, lines 4-6 ("conserved domain" refers to a polypeptide fragment that is more conserved at a sequence level than other fragments). The specification at page 10, lines 12-26 and page 10, lines 13-14 also notes conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for which residues may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority.

Support for new claim 55 is provided by, for example, page 5, line 9, page 6, lines 28-31, and page 12, line 27 through page 17, line 23.

Support for new claim 56 is provided by claim 12 as filed and on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 57 is provided by claim 12 as filed and on page 13, lines 15-19.

Support for new claim 58 is provided by claim 13 as filed and on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 59 is provided on, for example, page 17, lines 10-11 (utility for increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, lines 32 through page 21, line 28 (those genes induced after

D

exposure to biotropic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claim 60 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

Support for new claim 61 can be found in claim 1 as filed and page 6, lines 29-31 (transgenic plant comprising recombinant polynucleotide altering disease resistance or tolerance), page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests), page 11, lines 30-31 (identification of relation between two nucleic acid molecules through hybridization under stringent conditions), and on page 12, lines 1-4 (nucleic acids hybridize under stringent conditions to a probe under wash conditions of 0.2 x SSC, 0.1% SDS at 65° C).

Support for new claim 62 is provided by, for example, page 5, line 9, and page 16, line 9 through page 17, line 24.

Support for new claim 63 is provided on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 64 is provided on page 13, lines 15-19.

Support for new claim 65 is provided on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 66 is provided on, for example, page 17, lines 10-11 (increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, lines 32 through page 21, line 28 (those genes induced after exposure to biotropic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claim 67 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the

specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

Support for new claim 68 is may be found in claim 5 as filed (method for altering the disease tolerance or resistance of a plant), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests), page 11, lines 30-31 (identification of relation between two nucleic acid molecules through hybridization under stringent conditions), and on page 12, lines 1-4 (nucleic acids hybridize under stringent conditions to a probe under wash conditions of 0.2 x SSC, 0.1% SDS at 65° C).

Support for new claim 69 is provided by, for example, page 5, line 9, page 6, lines 28-31, and page 16, line 9 through page 17, line 24.

Support for new claim 70 is provided on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 71 is provided on page 13, lines 15-19.

Support for new claim 72 is provided on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 73 is provided by claim 5 as filed (method for altering the disease tolerance or resistance of a plant) and on, for example, page 17, lines 10-11 (increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, lines 32 through page 21, line 28 (those genes induced after exposure to biotropic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claim 74 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

Support for new claim 75 may be found at, for example, page 17, lines 10-11 (utility for increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines

10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, line 32 through page 21, line 28 (those genes induced after exposure to biotropic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance). The specification at page 10, lines 13-14 also notes conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for residues that may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority.

Support for new claim 76 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

RESPONSE TO REJECTIONS

Response to Rejection under 35 U.S.C. § 112, second paragraph

Claims 17-36 were rejected under 35 U.S.C. § 112, second paragraph, as they allegedly failed to particularly point out and distinctly claim the subject matter regarded as the invention. Applicants respectfully disagree.

Each of claims 17-36 have been canceled without prejudice or disclaimer and not for reasons of patentability. However, applicants submit that one of skill in the art would understand the metes and bounds of these claims. For example, the "conserved domain" of a transcription factor is a well-known concept in the field. Furthermore, the specification and the priority document specifically lists the conserved domains for a large number of transcription factors (*see* Table 1). Furthermore, that the conserved domains of transcription factors can be localization domains, activation domains, and DNA binding domains, for example, is also well known in the art. Applicants included a number of citations to references discussing the transcription factor families (*see* page 8, for example), and these and other references demonstrate that one of skill in

the art would understand the metes and bounds of each of the terms noted by the Examiner at page 5 of Paper No. 19.

The new claims 37-76 for the presently claimed transcription factors refer to only SEQ ID NO: 17 or 18 and are directed to a specific, defined conserved domain of residues 145-213 of SEQ ID NO: 18 which is supported by the specification as filed (see Figure 1a). That the conserved domains of transcription factors can be localization domains, activation domains, and DNA binding domains, for example, is also well known in the art. Applicants included a number of citations to references discussing the transcription factor families (*see* page 8, for example), and these and other references demonstrate that one of skill in the art would understand the metes and bounds of each of the terms noted by the Examiner at page 5 of Paper No. 19.

Applicants also submit that it would be abundantly clear to one of skill in the art what is meant by the term "conserved domain" in the present context. For example, the "conserved domain" of a transcription factor is a well-known concept in the field. See for example, Riechmann and Meyerowitz (1998), which is incorporated by reference (page 7, line 37 to page 8, line 1, and page 26, lines 29-30; submitted by Applicants in Information Disclosure Statement 11th October, 2002), and which defines conserved AP2 domains, with reference to, for example, the "conserved AP2 DNA-binding domain" (page 634, column 2, line 6); and "the conserved AP2 domain" (page 641, column 2, line 13). One of skill in the art would also be aware that the term "conserved domain" is often used in the art. For example, a search of issued U.S. patents yielded a sizeable number of references that make use of the term "conserved domain" (from 1996-2002 a search of ("conserved domain" AND sequence) revealed 338 patents that use the term in the present context).

The claims that were directed to "freezing, or nutrient, or pathogen stress" have been canceled, and the new claims are directed to "disease tolerance or resistance" or "tolerance to fungal disease", which are art-recognized terms and are defined in the specification; for example, page 7, lines 12-18, states: "[o]f particular interest are traits relating to increased disease resistance or tolerance of a plant, such as alterations in cell wall composition, trichome number or structure, callose induction, phytoalexin induction, alterations in the cell death response or the

like. These transgenic plants may be more resistant to biotrophic or necrotrophic pathogens such as a fungus, bacterium, mollicute, virus, nematode, a parasitic higher plant or the like and associated diseases. Another desirable phenotype is a change in the overall gene expression pattern of the plant in response to disease"). Page 17, lines 14-23 list pathogens and pests to which the plants can become tolerant.

Regarding the rejection based on improper Markush format, Applicant's believe the new claims avoid the rejection. The same may be said for the rejection based on indefiniteness of claims 22, 27 and 33, which referred to amino acid sequences, and claims that were dependent on indefinite claims.

The new claims make no reference to "DNA-binding domains." Therefore, even though one of skill in the art would clearly understand this term and its meaning, applicants need not address this reason for the rejection here.

Applicants submit that the new claims satisfy the requirements of 35 U.S.C. § 112, second paragraph, and respectfully request that this rejection be withdrawn.

Response to Rejection under 35 U.S.C. § 112, first paragraph

Claims 17-36 were rejected under 35 U.S.C. § 112, first paragraph, as the specification allegedly fails to reasonably convey that Applicants had possession of the invention. Applicants respectfully disagree with this rejection.

At page 6 of Paper No. 19, the Examiner characterizes the transcription factor of SEQ ID NO: 17 and 18 as "putative." Applicants respectfully but strongly disagree, as all the available and pertinent evidence indicates that SEQ ID NO: 17 encodes, and SEQ ID NO: 18 represents, a transcription factor, for at the very least the following reasons.

Applicants respectfully submit that the Examiner may not fully appreciate the PTO Written Description Examination Guidelines as they apply here. The USPTO Written Description Examination Guidelines state that:

> (t)he written description requirement for a claimed genus may satisfied through sufficient description of a representative number of species by actual reduction to practice ..., reduction to drawings ..., or by disclosure of relevant, identifying characteristics, i.e.,

D

> structure or other physical and/or chemical properties, by functional characteristics coupled with a known or disclosed correlation between function and structure, or by a combination of such identifying characteristics, sufficient to show the applicant was in possession of the claimed genus. (Federal Register, Jan. 5, 2001, Vol. 66(4) page 1106, II § 3 (a)(2).)

As shown in the attached report (Appendix F) and the Heard Declaration (Appendix C), Applicants correctly predicted the function of the sequence G28 (SEQ ID NO: 17; encoding SEQ ID NO: 18) based upon the identification of the sequence as an AP2 transcription factor and upon results from microarray experiments as disclosed in the instant application, and therefore Applicants characterized the sequence (G28) structurally and functionally. This is reflected by the present specification, which asserts in no uncertain terms that G28 is a transcription factor. See for example, page 4, line 30 through page 5, line 10: The transcription factor sequence may comprise a whole coding sequence or a fragment or domain of a coding sequence. A "fragment or domain, as referred to polypeptides, may be a portion of a polypeptide which performs at least one biological function of the intact polypeptide in substantially the same manner or to a similar extent as does the intact polypeptide. ... Exemplary polynucleotides or polypeptides comprise a sequence provided in the Sequence Listing as ... SEQ ID No.17 (G28), SEQ ID No.18 (G28 protein)." Applicants also clearly assert that G28 is a transcription factor in the specification at Figure 1a, because they assign a transcription factor family name (AP2) to the sequence.

Furthermore, as Fujimoto *et al.* ((2000) Plant Cell 12: 393-404, attached) point out, G28 (AtERF1; GenBank accession number AB008103) and related molecules are "members of a novel family of transcription factors that are specific to plants" (Abstract, lines 1-2). AtERF1 expression was also induced by ethylene and wounding, but not by other abiotic stresses such as cold, salinity, or drought (Fujimoto *et al.*, (2000), *supra*). Biotic stresses were apparently not tested. Since Fujimoto *et al.* published data identifying G28 as a transcription factor, they provide further evidence that Applicants were correct when they identified G28 as a transcription factor.

Furthermore, as Riechmann and Meyerowitz (1998) *supra*, point out "[f]ive amino acid residues are absolutely conserved among all of the AP2 domain sequences...[t]he conservation of those residues suggests their importance for the structure/function of the AP2 domain") p. 635,

D

column 2, lines 41-47). The sequence of G28 (SEQ ID NO: 18) with the five absolutely conserved residues shown in boldface in the underlined conserved domain below, are:

MSMTADSQSDYAFLESIRRHLLGESEPILSESTASSVTQSCVTGQSIKPVYGRNPSFSKLYPCFTESWGDL
PLKENDSEDMLVYGILNDAFHGGWEPSSSSSDEDRSSFPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTA
AKGKHYR**G**VRQRPW**G**KFAAEIRDPAKNGARVWLGTFETAEDAALAY**D**RAAFRMR**G**SRALL**N**FPLRVNSG
EPDPVRIKSKRSSFSSSNENGAPKKRRTVAAGGGMDKGLTVKCEVVEVARGDRLLVL

As noted in the Riechmann reference on page 635, column 2, lines 3-7, "[a] large number of sequences encoding AP2/EREBP proteins are already present in the databases and can be ***easily identified*** through BLAST searches due to the conservation of the AP2 domain, which, as mentioned above, is the only region conserved among all the proteins of the family (*emphasis* added)." Thus, one of skill in the art would not doubt that G28, having an AP2 conserved domain, and possessing the prerequisite five amino acid residues that are conserved among all of the AP2 domain sequences, is a member of the AP2 family of transcription factors.

Finally, and perhaps most significantly, Applicants have demonstrated the function of G28 as a transcription factor. In the declaration under 37 C.F.R. § 1.132 from Dr. Jacqueline Heard (Appendix C), who is a listed inventor of the instant application, Dr. Heard submits results of experiments performed prior to the date of filing the instant application. The results are presented as tables which show that Applicants had identified numerous polynucleotide sequences which endogenous mRNA levels were up-regulated by pathogens and/or methyl jasmonate, including G28 (SEQ ID NO: 17), G4 (SEQ ID NO: 12), G19 (SEQ ID NO: 22), G1006 (SEQ ID NO: 34), and G25 (SEQ ID NO: 82). Dr Heard also submits an internal report which discloses the effect of overexpressing G28 (SEQ ID NO: 17) in a transgenic plant. The report shows that Applicants' SEQ ID NO: 17 (and hence the encoded SEQ ID NO: 18) confers tolerance to fungal pathogens, just as the G19 (SEQ ID NO: 22) sequence noted above does.

Thus, the knowledge in the art, the above-described disclosures from the specification defining G28 as a transcription factor, coupled with the data submitted, all point to the same conclusion: G28 is a transcription factor.

With regard to the statement in Paper 19 that "Applicant's definition of a 'conserved domain' appears to be relative to SEQ ID NO: 18," Applicants note that the new claims are now directed to transcription factors comprising a conserved domain of a plant AP2 transcription factor wherein the conserved domain comprises one of an amino acid sequence of residues 145-213 of SEQ ID NO: 18, or an amino acid sequence having at least 84% identity to residues 145-213 of SEQ ID NO: 18, or an amino acid sequence of residues 145-213 of SEQ ID NO: 18 having one or more conservative substitutions, deletions, or insertions.

Applicants clearly identified amino acid residues 145-213 of SEQ ID NO: 18 at Figure 1a. The specification on page 11, lines 21-22 points out that "transcription factors that are homologs of the disclosed sequences will typically share at least 40% amino acid sequence identity," and page 6, lines 4-6: "[a] 'conserved domain' refers to a polynucleotide or polypeptide fragment that is more conserved at a sequence level than other fragments when the polynucleotide or polypeptide is compared with homologous genes or proteins from other plants."

The specification at page 10, lines 13-14 also notes the conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Support for which residues may be substituted and which residues they may be substituted with, is further disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814 (Appendix D), from which the instant application claims priority. In addition, fragments of domains of transcription factor sequences are specifically disclosed at page 4, lines 32-37, and in Figure 1.

Applicants also note that SEQ ID NO: 22 (G19) is disclosed in the specification as being significantly induced upon infection by the fungal pathogen *Erysiphe orontii* (*see* specification at page 25, lines 4-9). As shown in the accompanying BLAST report (Appendix F), G19 and G22 share 42% identity (note that the specification discloses that transcription factors that are homologs will typically share at least 40% amino acid sequence identity (page 11, lines 21-22). The conserved domain of G28 (SEQ ID NO: 18) shows even greater sequence identity to G19 (SEQ ID NO: 22), namely 66% identity (see attached BLAST report, page 11; handwritten notes regarding % identity). Thus, the BLAST report supports the new claims directed to an amino acid sequence having at least 42% identity to SEQ ID NO: 18. The use of identity ranges is well

known in the art and is specifically supported by the specification at, for example, page 11, lines 21-23. Thus, one of skill in the art would recognize the ability of the G19 sequence (SEQ ID NO: 22) and the closely related G28 sequence (SEQ ID NO: 18) to produce transgenic plants with disease tolerance or resistance, as is documented by Dr. Heard (see attached Declaration under 37 CFR 1.132, paragraphs 8 and 9 in particular).

The BLAST report also shows that conserved domains of G4 (SEQ ID NO: 12) with 62% identity to SEQ ID NO: 18, and G1006 (SEQ ID NO: 34) with 97% identity to SEQ ID NO: 18, and G25 (SEQ ID NO: 82) with 69% identity, show substantial sequence identity to the conserved domain of G28 (SEQ ID NO: 18) (*see* handwritten notes regarding % identity provided by the Applicants, particularly at pages 4, 10, and 13, respectively). The BLAST analysis method is fully disclosed in the specification at page 11, lines 3-20; at page 19, lines 2-8; and at page 25, lines 32-37. In addition, the specification discloses the function and utility of these polynucleotide and polypeptide sequences (see specification at pages 1 through 3, and pages 5 through 8, where the different polynucleotide and polypeptide sequences and their use are disclosed). Accordingly, Applicants' specification, in light of this evidence, clearly demonstrates the range of polynucleotide sequences that could be used in the new methods and transgenic plants.

As one of ordinary skill in the art recognizes, conserved domains may be identified as regions or domains of identity to a specific consensus sequence (see, for example, Riechmann et al., (2000) *Science* 290: 2105-2110; reference attached). Thus, by using alignment methods well known in the art, the conserved domains of the AP2 (APETALA2) domain transcription factor, the function of the presently claimed transcription factors may be determined family. Please see, for example, Riechmann and Meyerowitz (1998) *Biol. Chem.* 379: 633-646, which is incorporated by reference at page 7, line 37, continuing to page 8, line 1 of the specification (submitted by Applicants in Information Disclosure Statement 11th October, 2002). Riechmann and Meyerowitz specifically note:

(1) "Five amino acid residues are absolutely conserved among all of the AP2 domain sequences...[t]he conservation of those residues suggests their importance for the structure/function of the AP2 domain") p. 635, column 2, lines 41-47);

(2) "[a] primary source of specificity will be found to have very related, if not identical, DNA-binding specificities, similar to what has been observed for other large families of transcription factors" p. 643, column 1, lines 12-16; and

(4) "[p]artial genetic redundancy for *AP2* function and/or postranscriptional mechanisms of control of AP2 activity may underlie" (p. 640, column 2, lines 32-35).

Thus, at the time the present application was filed, one of skill in the art would recognize that portions of the AP2 domain are *absolutely conserved*, which strongly suggests their importance in imparting function and functional overlap, and that AP2 transcription factors are distinguishable by differential DNA-binding specificities, which are likely to be related, if not identical, within the family, and that there is genetic redundancy for AP2 function. One of skill in the art would instantly recognize that data obtained with G19-overexpressing plants would lead to the prediction that similar results would be obtained with plants overexpressing a transcription factor gene with a highly similar AP2 domain, as indicated in the present application.

Based on these published assertions, one of skill in the art would clearly have recognized from the knowledge available in the art, and from Applicants' specification and Figures, that G28 is a transcription factor and has the "only region conserved among all the proteins of the family" (see reference to Riechmann (1998), above), where conserved domains are listed for numerous sequences including SEQ ID NO: 18. Accordingly, Applicants' adequately demonstrated that they possessed sequences comprising a conserved domain as claimed. One of skill in the art could clearly perform the BLAST searches in order to print out and physically possess the sequences having at least 84% identity to the conserved domain of SEQ ID NO: 18, as recited in the new claims. In view of these facts, the art-recognized use of the terms "transcription factor" and "conserved domain," and the definition of a conserved AP2 domain (see above), Applicants submit that the new claims satisfy the written description requirement of 35 U.S.C. § 112, first paragraph.

To further demonstrate that the functional characteristics recited in the new claims would have been understood to be present with the recited polynucleotides, or easily identified in

D

transgenic plants comprising these polynucleotides, Applicants also enclose a copy of Table 1 of U.S. provisional application serial no. 60/125,814, at page 46 (Appendix D). The present application claims priority to this provisional application. Table 1 shows amino acid residues of a polypeptide sequence that may be substituted with other amino acids and are restricted to conservative substitutions.

Accordingly, one of skill in the art would not doubt that Applicants have adequately described the recited polynucleotides, specified a novel and useful transcription factor function of G28 and other AP2 family transcription factors in plants, and demonstrated that they can be made and used to enhance plant disease tolerance, for example, or that transgenic plants as claimed can be made and used. Applicants respectfully request reconsideration and withdrawal of the rejection.

Accordingly, applicants submit that the new claims satisfy the written description requirement of 35 U.S.C. § 112, first paragraph.

Claims 17-36 were further rejected under 35 U.S.C. § 112, first paragraph, as the specification allegedly failed to describe the subject matter of the claims in such a way as to enable one of skill in the art to make and use it. Applicants respectfully disagree.

The Examiner asserts that Applicants' specification fails to teach how SEQ ID NO: 18 can enhance a plant tolerance to stresses from pathogens, disease, freezing, etc. (*see* page 8 of Paper No: 19). The new claims are not directed to disease and freezing stress. As noted above, Dr. Heard has submitted that Applicants have in their possession data showing that SEQ ID NO: 17 confers pathogen or disease tolerance (*see* attached Declaration Appendix C, at exhibit H). One of skill in the art would not doubt these data. Applicants respectfully request reconsideration.

The Examiner has also asserted that the Quattrocchio document implies that undue experimentation would be required to produce a desired phenotype (*see* page 9 of Paper No. 19). First, the experiments disclosed in the specification at page 20, lines 11-16, and at page 24, lines 11-37 and continued on page 25, lines 1-2, show how one of skill in the art can determine the

plant traits as recited in the claims. These are the same types of experiments used to produce the results in the report for the G28 sequence.

Second, a careful reading of Quattrocchio shows that it does not support the undue experimentation position. Quattrocchio discusses the Petunia an2 gene, encoding a MYB-domain containing protein and that it had been suggested it could be the ortholog of the *Zea mays* c1 gene (Quattrocchio, page 476, first column, second paragraph, lines 2-5). The an2 gene had previously been shown to regulate anthocyanin biosynthesis (*see* Quattrocchio, page 478, first column, lines 1-2). Quattrocchio further states that "[t]he evolutionary model ***predicts*** that expression of c1 and an r gene in *P. hybrida* induces dfrA but not chsA, ***which was confirmed*** ..." and that "[t]he model also ***predicts*** that AN2 (the protein encoded by the an2 gene) should activate both early and late anthocyanin genes in *Z. mays*. ***This is precisely what we observed experimentally***" (*see* Quattrocchio, page 486, first column, first paragraph, lines 5-10; *emphasis* added). Thus, Quattrocchio shows that the MYB-encoding polynucleotide/genes an2 and c1 are functionally interchangeable between different plant species and that the same pathway of genes is controlled by the orthologous MYB-domain containing protein.

This interchangeability is further supported by Quattrocchio's statement that "[o]ur data for *P. hybrida* show that neither the *Z. mays* regulators c1 and lc ... nor an2 and jaf13 from *P. hybrida* can activate the early flavonoid genes ...." Therefore, the orthologous MYB-encoding polynucleotide sequences do not activate transcription inappropriately under the experimental conditions observed (*see* Quattrocchio, page 486, first column, fourth paragraph, lines 3-5 and continued in second column, line 1).

Nowhere does Quattrocchio suggest that the effect of transforming a plant with a heterologous "MYB" encoding polynucleotide "is unpredictable." In fact, Quattrocchio strongly suggests the very opposite, that transforming a plant with a heterologous MYB-encoding polynucleotide results in a predictable effect on the **same** pathway.

Also, Dr. Heard addresses the understanding one of skill in the art has for the contents of Quattrocchhio in her declaration (*see* page 4, in particular). Dr. Heard explains that the Quattrocchio document actually supports the predictable nature of using transcription factors to produce predictable effects on the same traits in plants. Further, the Duggleby (of record) and

Quattrocchio documents together, as discussed at pages 3-4 of the declaration, provide additional evidence of the ability of one of skill in the art to use the sequences recited in the claims. These statements demonstrate that the conclusions the Patent Office asserts from the Quattrocchio document are incorrect.

Accordingly, Applicants respectfully request reconsideration of the Quattrocchio document and this rejection.

Applicants also submit that a *prima facie* case of lack of enablement has not been made.

The remaining arguments are directed to the "homologous window sequence" recitation. Applicants note that the new claims do not recite the "homologous window sequence" discussed in this rejection. Accordingly, Applicants respectfully request reconsideration and that the Examiner withdraw the rejection.

Response to Rejection under 35 U.S.C. § 102

Claims 17-36 were rejected under 35 U.S.C. § 102(b) as being anticipated by Martin *et al.* (WO 97/47183).

Applicants note that the new claims do not recite a "homologous window sequence" or the "at least 6 consecutive amino acids" discussed in this rejection. Furthermore, the SEQ ID NO: 4 of Martin does not display the sequence identity to amino acid residues 145-213 of SEQ ID NO: 18 now recited in the claims. Accordingly, Applicants believe there is no anticipation of the present claims by the Martin reference.

Applicants respectfully request reconsideration and withdrawal of this rejection.

Provisional Double Patenting Rejection

The Examiner has rejected Claims 17-36 under the judicially created doctrine of obviousness-type double patenting as being unpatentable over copending applications 09/713,994 and 09/934,455.

As this is a provisional rejection, Applicants will address this rejection when claims of this or the copending application are allowed.

Applicants have requested an extension of time of two months in order to respond to the outstanding Office Action. No additional extension of time fees or requests for extension of time, or any other fees or petitions, are believed to be necessary to enter and consider this paper. If, however, any petitions or extensions of time are required or any fees are due in order to enter or consider this paper or enter or consider any paper accompanying this paper, including fees for net addition of claims, or in order to keep this application pending, Applicants hereby request any extensions or petitions necessary and the Commissioner is hereby authorized to charge Deposit Account No. 50-1129 for any fees.

Respectfully submitted,

WILEY REIN & FIELDING LLP

Date: October 29, 2002

By: _____
David J. Kulik
Registration No. 36,576

Enclosures: Appendix A (Marked-up version of amended Specification pages – 2 pages)
Appendix B (Gene Expression data; G1006 and G25 – 1 page)
Appendix C (executed declaration of Dr. Heard, pages 1-5 and Exhibits A-G)
Appendix D (pages 45 and 46 of priority application 60/125,814)
Appendix E (2 page conserved domain sequence comparison data for G28, G4, G19, G25, and G1006)
Appendix F (14 page BLAST report for G28, with handwritten notes)\
and
Fujimoto *et al.*, The Plant Cell 12:393-404 (2000)
Riechmann *et al.*, Science 290:2105-2110 (2000)

Wiley Rein & Fielding LLP
Attn: Patent Administration
1776 K Street, N.W.
Washington, D.C. 20006
202.719.7000 (Telephone)
202.719.7049 (Facsimile)

D

APPENDIX A

D

APPENDIX A
Marked-up Version of Amended Specification at pages 7-8

These transcription factors can be used to modulate a plant's response to disease. The plant transcription factors may belong to one of the following transcription factor families: the AP2 (APETALA2) domain transcription factor family (Riechmann and Meyerowitz (1998) **[J.]** Biol. Chem. 379:633-646); the MYB transcription factor family (Martin and Paz-Ares, (1997) Trends Genet. 13:67-73); the MADS domain transcription factor family (Riechmann and Meyerowitz (1997) **[J.]** Biol. Chem. 378:1079-1101); the WRKY protein family (Ishiguro and Nakamura (1994) Mol. Gen. Genet. 244:563-571); the ankyrin-repeat protein family (Zhang et al. (1992) Plant Cell 4:1575-1588); the zinc finger protein (Z) family (Klug and Schwabe (1995) FASEB J. 9: 597-604); the homeobox (HB) protein family (**Buerglin, in:** Duboule (1994) Guidebook to the Homeobox Genes, Oxford University Press, **pp. 27-71**); the CAAT-element binding proteins (Forsburg and Guarente (1989) Genes Dev. 3:1166-1178); the squamosa promoter binding proteins (SPB) (Klein et al. (1996) Mol. Gen. Genet. 1996 250:7-16); the NAM protein family (Souer et al. (1996) Cell 85:159-170); the IAA/AUX proteins (Rouse et al. (1998) Science 279:1371–1373); the HLH/MYC protein family (Littlewood et al. (1994) Prot. Profile 1:639-709); the DNA-binding protein (DBP) family (Tucker et al. (1994) EMBO J. 13:2994-3002); the bZIP family of transcription factors (Foster et al. (1994) FASEB J. 8:192-200); the Box P-binding protein (the BPF-1) family (da Costa e Silva et al. (1993) Plant J. 4:125-135); the high mobility group (HMG) family (Bustin and Reeves (1996) Prog. Nucl. Acids Res. Mol. Biol. 54:35-100); the scarecrow (SCR) family (Di Laurenzio et al. (1996) Cell 86:423-433); the GF14 family (Wu et al. (1997) Plant Physiol. 114:1421-1431); the polycomb (PCOMB) family (Kennison (1995) Annu. Rev. Genet. 29:289-303); the teosinte branched (TEO) family (Luo et al. (1996) Nature 383:794-799; the ABI3 family (Giraudat et al. (1992) Plant Cell 4:1251-1261); the triple helix (TH) family (Dehesh et al. (1990) Science 250:1397-1399); the EIL family (Chao et al. (1997) Cell 89:1133-44); the AT-HOOK family (Reeves and Nissen (1990) Journal of Biological Chemistry 265:8573-8582); the S1FA family (Zhou et al. (1995) Nucleic Acids Res. 23:1165-1169); the bZIPT2 family (Lu and Ferl (1995) Plant Physiol. 109:723); the YABBY

D

family (Bowman et al. (1999) Development 126:2387-96); the PAZ family (Bohmert et al. (1998) EMBO J. 17:170-80); a family of miscellaneous (MISC) transcription factors including the DPBF family (Kim et al. (1997) Plant J. 11:1237-1251) and the SPF1 family (Ishiguro and Nakamura (1994) Mol. Gen. Genet. 244:563-571); the golden (GLD) family (Hall et al. (1998) Plant Cell 10:925-936).

APPENDIX B

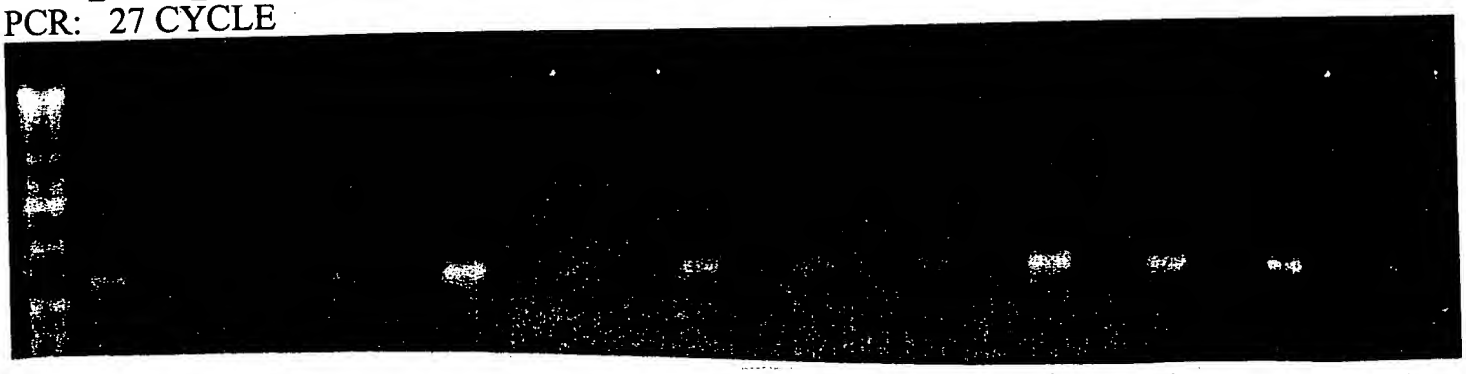
Appendix B

G1006

Gene Expression in Different Conditions

Auxin (soil)	ABA (soil)	Cold (soil)	Heat (soil)	Salt (soil)	Drought (soil)	Osmotic (soil)	Control (soil)	Erysiphe (soil)	Fusarium (plate)	SA (plate)	Control (plate)
L	L	L	L	L	L	L	L	L	L	L	L

PCR: 27 CYCLE



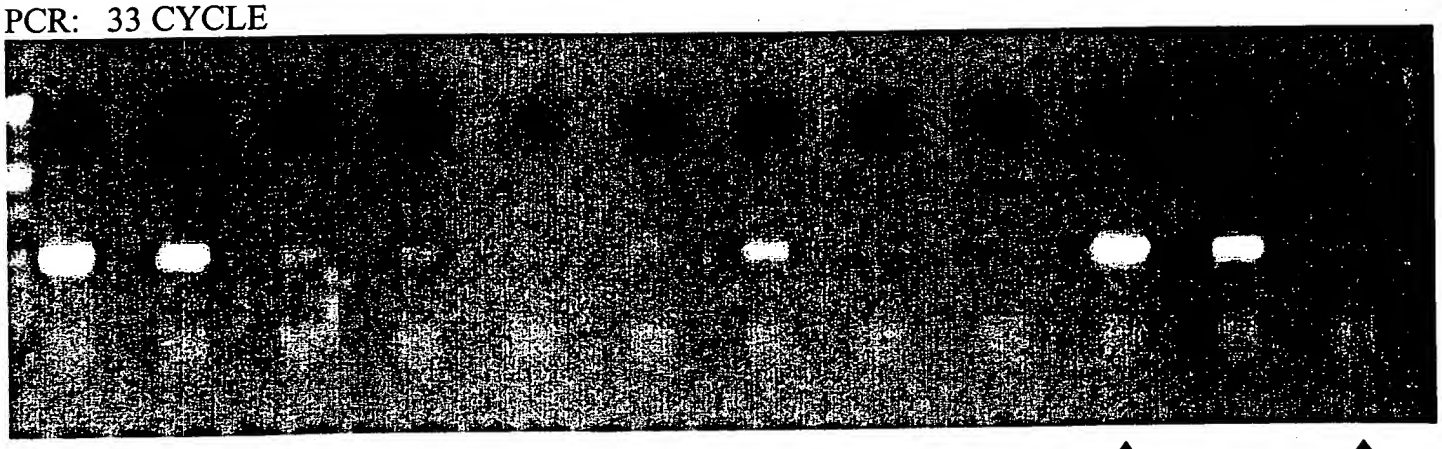
Control (in soil) Eryisphe (in soil) Fusarium (on plate) Control (on plate)

G25

Gene Expression in Different Conditions

Auxin (soil)	ABA (soil)	Cold (soil)	Heat (soil)	Salt (soil)	Drought (soil)	Osmotic (soil)	Control (soil)	Erysiphe (soil)	Fusarium (plate)	SA (plate)	Control (plate)
H	H	L	L	L	L	L	L	L	H	H	L

PCR: 33 CYCLE



Fusarium (on plate) Control (on plate)

APPENDIX C

Executed Declaration of Jacqueline E. Heard Under 37 CFR 1.132

Exhibit A: *Fusarium* treatment after 24 hours

Exhibit B: *Fusarium* treatment after 48 hours

Exhibit C: *Erysiphe* treatment after 7 days

Exhibit D: Methyl Jasmonate treatment after 24 hours

Exhibit E: *Botrytis* treatment after 12 hours

Exhibit F: *Erysiphe* treatment

Exhibit G: Methyl Jasmonate treatment

Exhibit H: Summary of Overexpressor G28, Family AP2

APPENDIX D

share at least 85%, 90% or 95% sequence identity. The homologs will also have substantially the same DNA binding specificity. At the nucleotide level, the sequences will typically share at least 40% nucleotide sequence identity, preferably at least 50%, 60%, 70% or 80% sequence identity, and more preferably 85%, 90%, 95% or 97% sequence identity.

Homologs from the same plant, different plant species or other organisms may be identified using sequence alignment methods and homology calculations, such as those described in Altschul et al. (1994) Nature Genetics 6: 119-129. For example, the NCBI Basic Local Alignment Search Tool (BLAST) (Altschul et al. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215:403-410). is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastp, tblastn and tblastx.

Substitutions, deletions and insertions intorduced in the DNA binding domain are also envisioned by this invention. Amino acid substitutions are typically of single residues; insertions usually will be on the order of about from 1 to 10 amino acid residues; and deletions will range about from 1 to 30 residues. Deletions or insertions preferably are made in adjacent pairs, i.e., a deletion of two residues or insertion of two residues. Substitutions, deletions, insertions or any combination thereof may be combined to arrive at a final construct. Obviously, the mutations that are made in the DNA encoding the protein must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure.

Substitutional variants are those in which at least one residue in the amino acid sequence has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the following Table 1 when it is desired to finely modulate the characteristics of the protein. Table 1 shows amino acids which may be substituted for an original amino acid in a protein and which are typically regarded as conservative substitutions.

Table 1

Original Residue	Conservative Substitutions
Ala	ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser
Gln	asn
Glu	asp
Gly	pro
His	asn; gln
Ile	leu, val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

An alternative indication is to show whethr two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5°C to 20°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Conditions for nucleic acid hybridization and calculation of stringencies can be found in Sambrook et al., Molecular Cloning. A Laboratory Manual, Ed. 2, Cold Spring Harbor Laboratory Press, New York (1989)) and Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes* Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993). Nucleic acid molecules that hybridize under stringent conditions will typically hybridize to a probe based on either the entire cDNA

APPENDIX E

Appendix E

Percent identity comparisons of conserved domains (underlined) of subject sequences of AP2 transcription factor family members as disclosed in specification USSN 09/533,029.

SEQ ID NO:18 (G28) (query sequence) Conserved domain = residues 145-213
MSMTADSQSDYAFLESIRRHLLGESEPILSESTASSVTQSCVTGQSIKPVYGRNPSFSKLYPCFTESWGDLPLKEND
SEDMLVYGILNDAFHGGWEPSSSSSDEDRSSFPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQR
PWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPDPVRIKSKRSSFSSSNEN
GAPKKRRTVAAGGGMDKGLTVKCEVVEVARGDRLLVL

BLAST comparisons of subject sequences disclosed in specification USSN 09/533,029

SEQ ID NO:12 (G4) Conserved domain = residues 121-188
MCGGAIISDFIPPPRSLRVTNEFIWPDLKNKVKASKKRSNKRSDFFDLDDDFEADFQGFKDDSAFDCEDDDDVFVNV
KPFVFTATTKPVASAFVSTVGSAYAKKTVESAEQAEKSSKRKRKNQYRGIRQRPWGKWAAEIRDPRKGSREWLGTFD
TAEEAARAYDAAARRIRGTKAKVNFPEEKNPSVVSQKRPSAKTNNLQKSVAKPNKSVTLVQQPTHLSQQYCNNSFDN
SFGDMSFMEEKPQMYNNQFGLTNSFDAGGNNGYQYFSSDQGSNSFDCSEFGWSDHGPKTPEISSMLVNNNEASFVEE
TNAAKKLKPNSDESDDLMAYLDNALWDTPLEVEAMLGADAGAVTQEEENPVELWSLDEINFMLEGDF

SEQ ID NO:12 (G4) (Amino Acid Sequence) (gf=1)
Length = 375

Score = 90.5 bits (223), Expect = 1e-20
Identities = 43/62 (69%), Positives = 52/62 (83%), Gaps = 1/62 (1%)

```
Query: 4    YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLR 63
            YRG+RQRPWGK+AAEIRDP K G+R WLGTF+TAE+AA AYD AA R+RG++A +NFP
Sbjct: 124  YRGIRQRPWGKWAAEIRDPRK-GSREWLGTFDTAEEAARAYDAAARRIRGTKAKVNFPEE 182

Query: 64   VN 65
             N
Sbjct: 183  KN 184
```

SEQ ID NO:22 (G19) Conserved domain = residues 76-145
MCGGAIISDYAPLVTKAKGRKLTAEELWSELDASAADDFWGFYSTSKLHPTNQVNVKEEAVKKEQATEPGKRRKRKN
VYRGIRKRPWGKWAAEIRDPRKGVRVWLGTFNTAEEAAMAYDVAAKQIRGDKAKLNFPDLHHPPPPNYTPPPSSPRS
TDQPPAKKVCVVSQSESELSQPSFPVECIGFGNGDEFQNLSYGFEPDYDLKQQISSLESFLELDGNTAEQPSQLDES
VSEVDMWMLDDVIASYE

SEQ ID NO:22 (G19) (Amino Acid Sequence) (gf=1)
Length = 248

Score = 89.7 bits (221), Expect = 2e-20
Identities = 42/58 (72%), Positives = 50/58 (85%), Gaps = 1/58 (1%)

```
Query: 4    YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 61
            YRG+R+RPWGK+AAEIRDP K G RVWLGTF TAE+AA+AYD AA ++RG +A LNFP
Sbjct: 79   YRGIRKRPWGKWAAEIRDPRK-GVRVWLGTFNTAEEAAMAYDVAAKQIRGDKAKLNFP 135
```

SEQ ID NO:82 (G25) Conserved domain = residues 47-114
MCGGAIISDFIWSKSESEPSQLGSVSSRKKRKPVSVSEERDGKRERKNLYRGIRQRPWGKWAAEIRDPSKGVRVWLG
TFKTADEAARAYDVAAIKIRGRKAKLNFPNTQVEEEADTKPGGNQNELISENQVESLSEDLMALEDYMRFYQIPVAD
DQSATDIGNLWSYQDSN

SEQ ID NO:82 (G25) (Amino Acid Sequence) (gf=1)
Length = 171

Score = 90.9 bits (224), Expect = 8e-21
Identities = 42/58 (72%), Positives = 51/58 (87%), Gaps = 1/58 (1%)

```
Query: 4     YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 61
              YRG+RQRPWGK+AAEIRDP+K G RVWLGTF+TA++AA AYD AA ++RG +A LNFP
Sbjct: 50    YRGIRQRPWGKWAAEIRDPSK-GVRVWLGTFKTADEAARAYDVAAIKIRGRKAKLNFP 106
```

SEQ ID NO:34 (G1006) Conserved domain = residues 114-182
MYGQCNIESDYALLESITRHLLGGGGENELRLNESTPSSCFTESWGGLPLKENDSEDMLVYGLLKDAFHFDTSSSDL
SCLFDFPAVKVEPTENFTAMEEKPKKAIPVTETAVKAKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAA
LAYDIAAFRMRGSRALLNFPLRVNSGEPDPVRITSKRSSSSSSSSSSTSSSENGKLKRRRKAENLTSEVVQVKCEV
GDETRVDELLVS

SEQ ID NO:34 (G1006) (Amino Acid Sequence) (gf=1)
Length = 243

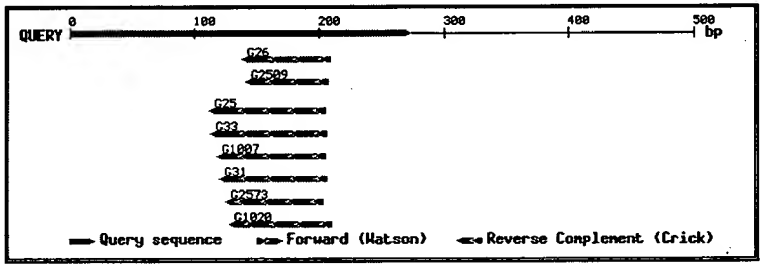
Score = 140 bits (353), Expect = 9e-36
Identities = 67/68 (98%), Positives = 67/68 (98%)

```
Query: 2     KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 61
              KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYD AAFRMRGSRALLNFP
Sbjct: 115   KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDIAAFRMRGSRALLNFP 174

Query: 62    LRVNSGEP 69
              LRVNSGEP
Sbjct: 175   LRVNSGEP 182
```

APPENDIX F

Summary of BLAST against PTFGENE dataset, minimum match cutoff 50%.



graphics by Shuai Weng, Stanford
The full results of this search are below.

Mendel BLAST query on *Arabidopsis* sequences

Full BLAST options and parameters, refer to the NCBI BLAST Documentation. Links to GenBank, EMBL, PIR, and SwissProt are shown in **bold** type; links to locations within this document are in normal type.

BLASTP 2.1.3 [Apr-1-2001]

References:

- Altschul, Stephen F., Gish, Warren, Miller, Webb, Myers, Eugene W., and David J. Lipman (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410. [PubMed]
- Gish, Warren, and David J. States (1993). Identification of protein coding regions by database similarity search. *Nature Genetics* 3:266-72. [PubMed]
- Gish, Warren (1994) unpublished. BLAST2 Documentation

Query Sequence: G28 (Length: 268)

Database: ptfgene 1566 sequences; 551,722 total letters

Sequences producing significant alignments:	(bits)	Value
G28 (Amino Acid Sequence) (gf=1)	535	e-154
G1006 (Amino Acid Sequence) (gf=1)	268	2e-73
G22 (Amino Acid Sequence) (gf=1)	161	2e-41
G1004 (Amino Acid Sequence) (gf=1)	127	6e-31
G29 (Amino Acid Sequence) (gf=1)	117	6e-28
G1266 (Amino Acid Sequence) (gf=1)	113	7e-27
G46 (Amino Acid Sequence) (gf=1)	110	4e-26

G43 (Amino Acid Sequence) (gf=1)	110	4e-26
G1791 (Amino Acid Sequence) (gf=1)	110	8e-26
G1792 (Amino Acid Sequence) (gf=1)	107	4e-25
G1841 (Amino Acid Sequence) (gf=1)	107	5e-25
G2512 (Amino Acid Sequence) (gf=1)	107	6e-25
G1419 (Amino Acid Sequence) (gf=1)	107	6e-25
G1411 (Amino Acid Sequence) (gf=1)	104	3e-24
G865 (Amino Acid Sequence) (gf=1)	104	3e-24
G30 (Amino Acid Sequence) (gf=1)	103	7e-24
G44 (Amino Acid Sequence) (gf=1)	103	9e-24
G1795 (Amino Acid Sequence) (gf=1)	100	5e-23
G1385 (Amino Acid Sequence) (gf=1)	100	6e-23
G1752 (Amino Acid Sequence) (gf=1)	100	8e-23
G1751 (Amino Acid Sequence) (gf=1)	100	8e-23
G2509 (Amino Acid Sequence) (gf=1)	99	1e-22
G26 (Amino Acid Sequence) (gf=1)	99	1e-22
G1750 (Amino Acid Sequence) (gf=1)	99	2e-22
G864 (Amino Acid Sequence) (gf=1)	99	2e-22
G4 (Amino Acid Sequence) (gf=1)	98	4e-22
G2130 (Amino Acid Sequence) (gf=1)	97	9e-22
G7 (Amino Acid Sequence) (gf=1)	97	9e-22
G45 (Amino Acid Sequence) (gf=1)	96	1e-21
G19 (Amino Acid Sequence) (gf=1)	96	1e-21
G14 (Amino Acid Sequence) (gf=1)	96	2e-21
G1008 (Amino Acid Sequence) (gf=1)	94	4e-21
G1421 (Amino Acid Sequence) (gf=1)	94	6e-21
G1794 (Amino Acid Sequence) (gf=1)	94	7e-21
G440 (Amino Acid Sequence) (gf=1)	93	1e-20
G31 (Amino Acid Sequence) (gf=1)	92	2e-20
G25 (Amino Acid Sequence) (gf=1)	92	2e-20
G1763 (Amino Acid Sequence) (gf=1)	91	4e-20
G1384 (Amino Acid Sequence) (gf=1)	91	4e-20
G974 (Amino Acid Sequence) (gf=1)	91	4e-20
G33 (Amino Acid Sequence) (gf=1)	91	5e-20
G1932 (Amino Acid Sequence) (gf=1)	91	6e-20
G1005 (Amino Acid Sequence) (gf=1)	91	6e-20
G439 (Amino Acid Sequence) (gf=1)	89	1e-19
G2135 (Amino Acid Sequence) (gf=1)	88	3e-19
G1380 (Amino Acid Sequence) (gf=1)	88	3e-19
G625 (Amino Acid Sequence) (gf=1)	88	3e-19
G1748 (Amino Acid Sequence) (gf=1)	87	5e-19
G1845 (Amino Acid Sequence) (gf=1)	87	7e-19
G2059 (Amino Acid Sequence) (gf=1)	86	1e-18
G1020 (Amino Acid Sequence) (gf=1)	86	1e-18
G6 (Amino Acid Sequence) (gf=1)	86	1e-18
G1796 (Amino Acid Sequence) (gf=1)	86	2e-18
G5 (Amino Acid Sequence) (gf=1)	86	2e-18
G2113 (Amino Acid Sequence) (gf=1)	85	3e-18
G2571 (Amino Acid Sequence) (gf=1)	85	3e-18
G36 (Amino Acid Sequence) (gf=1)	84	4e-18
G32 (Amino Acid Sequence) (gf=1)	82	2e-17
G35 (Amino Acid Sequence) (gf=1)	82	2e-17
G1415 (Amino Acid Sequence) (gf=1)	81	4e-17
G23 (Amino Acid Sequence) (gf=1)	80	8e-17
G12 (Amino Acid Sequence) (gf=1)	80	8e-17
G1141 (Amino Acid Sequence) (gf=1)	80	1e-16
G2138 (Amino Acid Sequence) (gf=1)	79	1e-16
G38 (Amino Acid Sequence) (gf=1)	79	1e-16
G975 (Amino Acid Sequence) (gf=1)	79	2e-16
G1800 (Amino Acid Sequence) (gf=1)	78	3e-16
G442 (Amino Acid Sequence) (gf=1)	76	1e-15
G1754 (Amino Acid Sequence) (gf=1)	76	2e-15
G1090 (Amino Acid Sequence) (gf=1)	76	2e-15
G1007 (Amino Acid Sequence) (gf=1)	76	2e-15
G2299 (Amino Acid Sequence) (gf=1)	75	2e-15

G16 (Amino Acid Sequence) (gf=1)	75	2e-15
G2134 (Amino Acid Sequence) (gf=1)	75	3e-15
G977 (Amino Acid Sequence) (gf=1)	75	3e-15
G24 (Amino Acid Sequence) (gf=1)	75	4e-15
G10 (Amino Acid Sequence) (gf=1)	75	4e-15
G1387 (Amino Acid Sequence) (gf=1)	74	6e-15
G1386 (Amino Acid Sequence) (gf=1)	74	6e-15
G1379 (Amino Acid Sequence) (gf=1)	73	1e-14
G441 (Amino Acid Sequence) (gf=1)	73	1e-14
G39 (Amino Acid Sequence) (gf=1)	73	1e-14
G2583 (Amino Acid Sequence) (gf=1)	73	1e-14
G1846 (Amino Acid Sequence) (gf=1)	73	1e-14
G1755 (Amino Acid Sequence) (gf=1)	72	2e-14
G1277 (Amino Acid Sequence) (gf=1)	72	2e-14
G21 (Amino Acid Sequence) (gf=1)	72	2e-14
G2573 (Amino Acid Sequence) (gf=1)	72	3e-14
G27 (Amino Acid Sequence) (gf=1)	72	3e-14
G1381 (Amino Acid Sequence) (gf=1)	71	4e-14
G2510 (Amino Acid Sequence) (gf=1)	71	5e-14
G2291 (Amino Acid Sequence) (gf=1)	70	7e-14
G2108 (Amino Acid Sequence) (gf=1)	70	1e-13
G3 (Amino Acid Sequence) (gf=1)	70	1e-13
G13 (Amino Acid Sequence) (gf=1)	69	2e-13
G2115 (Amino Acid Sequence) (gf=1)	68	3e-13
G869 (Amino Acid Sequence) (gf=1)	68	4e-13
G40 (Amino Acid Sequence) (gf=1)	67	6e-13
G41 (Amino Acid Sequence) (gf=1)	67	1e-12
G2067 (Amino Acid Sequence) (gf=1)	66	1e-12

>G28 (Amino Acid Sequence) (gf=1)
Length = 268

Score = 535 bits (1378), Expect = e-154
Identities = 268/268 (100%), Positives = 268/268 (100%)

```
Query: 1   MSMTADSQSDYAFLESIRRHLLGESEPILSESTASSVTQSCVTGQSIKPVYGRNPSFSKL 60
           MSMTADSQSDYAFLESIRRHLLGESEPILSESTASSVTQSCVTGQSIKPVYGRNPSFSKL
Sbjct: 1   MSMTADSQSDYAFLESIRRHLLGESEPILSESTASSVTQSCVTGQSIKPVYGRNPSFSKL 60

Query: 61  YPCFTESWGDLPLKENDSEDMLVYGILNDAFHGGWEPSSSSSDEDRSSFPSVKIETPESF 120
           YPCFTESWGDLPLKENDSEDMLVYGILNDAFHGGWEPSSSSSDEDRSSFPSVKIETPESF
Sbjct: 61  YPCFTESWGDLPLKENDSEDMLVYGILNDAFHGGWEPSSSSSDEDRSSFPSVKIETPESF 120

Query: 121 AAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFET 180
           AAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFET
Sbjct: 121 AAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFET 180

Query: 181 AEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPDPVRIKSKRSSFSSSNENGAPKKRRTV 240
           AEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPDPVRIKSKRSSFSSSNENGAPKKRRTV
Sbjct: 181 AEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPDPVRIKSKRSSFSSSNENGAPKKRRTV 240

Query: 241 AAGGGMDKGLTVKCEVVEVARGDRLLVL 268
           AAGGGMDKGLTVKCEVVEVARGDRLLVL
Sbjct: 241 AAGGGMDKGLTVKCEVVEVARGDRLLVL 268
```

>G1006 (Amino Acid Sequence) (gf=1)
Length = 243

Score = 268 bits (685), Expect = 2e-73
Identities = 163/271 (60%), Positives = 180/271 (66%), Gaps = 47/271 (17%)

```
Query: 8   QSDYAFLESIRRHLLG---ESEPILSESTASSVTQSCVTGQSIKPVYGRNPSFSKLYPCF 64
           +SDYA LESI RHLLG   E+E  L+EST SS                          CF
Sbjct: 8   ESDYALLESITRHLLGGGGENELRLNESTPSS-------------------------CF 41
```



```
Query: 65  TESWGDLPLKENDSEDMLVYGILNDAFHGGWEPSSSSSDEDRSSFPSVKIETPESFAAVD 124
           TESWG LPLKENDSEDMLVYG+L DAFH      +SSS        FP+VK+E  E+F A++
Sbjct: 42  TESWGGLPLKENDSEDMLVYGLLKDAFHFD---TSSSDLSCLFDFPAVKVEPTENFTAME 98

Query: 125 SVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDA 184
             P   +K  PV+    TA K KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDA
Sbjct: 99  EKP---KKAIPVTE--TAVKKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDA 153

Query: 185 ALAYDRAAFRMRGSRALLNFPLRVNSGEPDPVRIKSKRSSFS--------SSNENGAPKK 236
           ALAYD AAFRMRGSRALLNFPLRVNSGEPDPVRI SKRSS S         SS+ENG  K+
Sbjct: 154 ALAYDIAAFRMRGSRALLNFPLRVNSGEPDPVRITSKRSSSSSSSSSSTSSSENGKLKR 213

Query: 237 RRTVAAGGGMDKGLTVKCEVVEVARGDRLLV 267
           RR   A      + + VKCEV +  R D LLV
Sbjct: 214 RR--KAENLTSEVVQVKCEVGDETRVDELLV 242
```

>G22 (Amino Acid Sequence) (gf=1)
Length = 226

Score = 161 bits (408), Expect = 2e-41
Identities = 87/175 (49%), Positives = 107/175 (60%), Gaps = 21/175 (12%)

```
Query: 50  VYGRNPSFSKLYPCPTESWGDLPLKENDSEDMLVYGILNDAFHGGWEPSSSSSDEDRSSF 109
           +Y RNPSFS +      ++W DLPL  +DS+DM +Y  L DA    GW PS
Sbjct: 15  MYFRNPSFSNVI--LNDNWSDLPLSVDDSQDMAIYNTLRDAVSSGWTPS----------V 62

Query: 110 PSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKN 169
           P V     E+          K   T    +     KG  YRGVR+RPWGKFAAEIRDP KN
Sbjct: 63  PPVTSPAEEN---------KPPATKASGSHAPRQKGMQYRGVRRRPWGKFAAEIRDPKKN 113

Query: 170 GARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPDPVRIKSKRSS 224
           GARVWLGT+ET EDAA+AYDRAAF++RGS+A LNFP  + S + +PVRI+ +R S
Sbjct: 114 GARVWLGTYETPEDAAVAYDRAAFQLRGSKAKLNFPHLIGSCKYEPVRIRPRRRS 168
```

>G1004 (Amino Acid Sequence) (gf=1)
Length = 300

Score = 127 bits (318), Expect = 6e-31
Identities = 91/227 (40%), Positives = 126/227 (55%), Gaps = 30/227 (13%)

```
Query: 17  IRRHLLGESEPILSES----TASSVTQSCVTGQSIKPVYGRNPSFSKLYPC-FTESWGDL 71
           I +HLL E+ P+ ++       +SS T+S     SI  ++G   S S   P  F+ES
Sbjct: 13  IEKHLLDEASPVATDPWMKHESSSATESSSDSSSI--IFGS--SSSSFAPIDFSESVCKP 68

Query: 72  PLKENDSEDMLVYGILNDAFHGGWEPSSSS-----SDEDRSSF----------PSVKIET 116
            + + D+   + +  L+  F    E S S      S+++++ F          P +KI
Sbjct: 69  EIIDLDTPRSMEF--LSIPFEFDSEVSVSDFDFKPSNQNQNQFEPELKSQIRKPPLKISL 126

Query: 117 PESFAAVDSVPVKKEKTSP-VSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWL 175
           P   A  + +    E T P V+  V+  + KHYRGVRQRPWGKFAAEIRDP K G+RVWL
Sbjct: 127 P---AKTEWIQFAAENTKPEVTKPVSEEEKKHYRGVRQRPWGKFAAEIRDPNKRGSRVWL 183

Query: 176 GTFETAEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPDPVRIKSKR 222
           GTF+TA +AA AYD AAFR+RGS+A+LNFPL V    +P     + KR
Sbjct: 184 GTFDTAIEAARAYDEAAFRLRGSKAILNFPLEVGKWKPRADEGEKKR 230
```

>G29 (Amino Acid Sequence) (gf=1)
Length = 282

Score = 117 bits (292), Expect = 6e-28
Identities = 65/145 (44%), Positives = 84/145 (57%), Gaps = 6/145 (4%)

67/69 = 97%

```
Query: 88  NDAFHGGWEPSSSSSDEDRSSFPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKH 147
           ND+F     P   +   +R   P +KI  P     +       +   PV   V A + +H
Sbjct: 80  NDSFTFQSNPPRVTVQSNRK--PPLKIAPPNRTKWIQFATGNPKPELPVPV-VAAEEKRH 136

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLR 207
           YRGVR RPWGKFAAEIRDP + G RVWLGTFETA +AA AYD+ AFR+RGS+A+LNFPL
Sbjct: 137 YRGVRMRPWGKFAAEIRDPTRRGTRVWLGTFETAIEAARAYDKEAFRLRGSKAILNFPLE 196

Query: 208 VNSGEPDPVRIKSKRSSFSSSNENG 232
           V+   P   R +  R  ++    +G
Sbjct: 197 VDKWNP---RAEDGRGLYNKRKRDG 218


>G1266 (Amino Acid Sequence) (gf=1)
         Length = 218

 Score =  113 bits (283), Expect = 7e-27
 Identities = 61/134 (45%), Positives = 79/134 (58%), Gaps = 31/134 (23%)

Query: 71  LPLKENDSEDMLVYGILNDAFHGGWEPSSSSSDEDRSSPPSVKIETPESFAAVDSVPVKK 130
           LP  ENDSE+M +YG++  +                              T +++   DS
Sbjct: 36  LPFNENDSEEMFLYGLIEQS------------------------TQQTYIDSDS----- 65

Query: 131 EKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDR 190
            +  P+ +  +    K YRGVR+RPWGKFAAEIRD  +NG RVWLGTFE+AE+AALAYD+
Sbjct: 66  -QDLPIKSVSSRKSEKSYRGVRRRPWGKFAAEIRDSTRNGIRVWLGTFESAEEAALAYDQ 124

Query: 191 AAFRMRGSRALLNF 204
           AAF MRGS A+LNF
Sbjct: 125 AAFSMRGSSAILNF 138


>G46 (Amino Acid Sequence) (gf=1)
         Length = 207

 Score =  110 bits (276), Expect = 4e-26
 Identities = 73/154 (47%), Positives = 89/154 (57%), Gaps = 16/154 (10%)

Query: 71  LPLKENDSEDMLVYGILNDAFHGGWEPSSSSSDEDRSSFPSVKIETPESFAAVDSVPVKK 130
           +P +E DS  +     +N+      E SSSSS E  SS  S   ET +S
Sbjct: 48  VPKQEPDSPVLDPDSFVNEFLQVEGESSSSSSPELNSS--SSTYETDQSV---------- 95

Query: 131 EKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDR 190
           +K       V A   +HYRGVR+RPWGKFAAEIRDPAK G+R+WLGTFE+   DAA AYD
Sbjct: 96  KKAERFEEEVDA---RHYRGVRRRPWGKFAAEIRDPAKKGSRIWLGTFESDVDAARAYDC 152

Query: 191 AAFRMRGSRALLNFPLRVNSGE-PDPVRIKSKRS 223
           AAF++RG +A+LNFPL     E P     K KRS
Sbjct: 153 AAFKLRGRKAVLNFPLDAGKYEAPANSGRKRKRS 186


>G43 (Amino Acid Sequence) (gf=1)
         Length = 201

 Score =  110 bits (276), Expect = 4e-26
 Identities = 62/128 (48%), Positives = 82/128 (63%), Gaps = 3/128 (2%)

Query: 96  EPSSSSSDEDRSSFPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRP 155
           EPSS   D D      +++E   S ++  +   + E  S          + +HYRGVR+RP
Sbjct: 55  EPSSPVLDPDSYVQEILQMEAESSSSSSTTTSPEVETVSNRKKTKRFEETRHYRGVRRRP 114

Query: 156 WGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPD- 214
           WGKFAAEIRDPAK G+R+WLGTFE+  DAA AYD AAF++RG +A+LNFPL  ++G+ D
Sbjct: 115 WGKFAAEIRDPAKKGSRIWLGTFESDIDAARAYDYAAFKLRGRKAVLNFPL--DAGKYDA 172
```



```
Query: 215 PVRIKSKR 222
           PV    KR
Sbjct: 173 PVNSCRKR 180


>G1791 (Amino Acid Sequence) (gf=1)
         Length = 139

 Score =  110 bits (274), Expect = 8e-26
 Identities = 58/105 (55%), Positives = 66/105 (62%), Gaps = 10/105 (9%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLR 207
           YRGVR+RPWGK+AAEIRD A++GARVWLGTF TAEDAA AYDRAAF MRG RA+LNFP
Sbjct: 14  YRGVRKRPWGKYAAEIRDSARHGARVWLGTFNTAEDAARAYDRAAFGMRGQRAILNFPHE 73

Query: 208 VNSGEPDPVRIKSKRSSFSSSNENGAPKKRRTVAAGGGMDKGLTV 252
           V   P          + S+EN           GGG D G  V
Sbjct: 74  YQMMKDGP----------NGSHENAVASSSSGYRGGGGGDDGREV 108


>G1792 (Amino Acid Sequence) (gf=1)
         Length = 139

 Score =  107 bits (268), Expect = 4e-25
 Identities = 55/92 (59%), Positives = 69/92 (74%), Gaps = 5/92 (5%)

Query: 144 KGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLN 203
           K   +RGVR+RPWGKFAAEIRDP++NGAR+WLGTFETAE+AA AYDRAAF +RG  A+LN
Sbjct: 16  KQARPRGVRRRPWGKFAAEIRDPSRNGARLWLGTFETAEEAARAYDRAAFNLRGHLAILN 75

Query: 204 FPLRVNSGEP--DPVRIKSKRSSFSSSNENGA 233
           FP   N   P  D   ++   +S SSS+ +G+
Sbjct: 76  FP---NEYYPRMDDYSLRPPYASSSSSSSGS 104


>G1841 (Amino Acid Sequence) (gf=1)
         Length = 229

 Score =  107 bits (267), Expect = 5e-25
 Identities = 66/139 (47%), Positives = 80/139 (57%), Gaps = 4/139 (2%)

Query: 94  GWEPSSSSSDEDRSSFPSVKIETPESFAAVDSV--PVKKEKTSPVSAAVTAAKGKHYRGV 151
           G   SSSS+   R           P SFA VDS   P++ E  S          K + YRGV
Sbjct: 31  GAHSSSSSAGHKRGWLGIDSAPIPSSFARVDSSHNPIE-ESMSKAFPEEAREKKRRYRGV 89

Query: 152 RQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLRVNSG 211
           RQRPWGK+AAEIRDP +  ARVWLGTF+TAE AA AYD AA R RG++A LNFP  V
Sbjct: 90  RQRPWGKWAAEIRDPHR-AARVWLGTFDTAEAAARAYDEAALRFRGNKAKLNFPEDVRIL 148

Query: 212 EPDPVRIKSKRSSFSSSNE 230
           EP P  ++S   + ++  E
Sbjct: 149 PPPPPLLRSPADTVANKAE 167


>G2512 (Amino Acid Sequence) (gf=1)
         Length = 244

 Score =  107 bits (266), Expect = 6e-25
 Identities = 64/143 (44%), Positives = 82/143 (56%), Gaps = 16/143 (11%)

Query: 95  WEPSSSSSDEDRSSFPSVKIETP-ESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQ 153
           WE S      D+S   S   +   + F A +S  +K+E      AA     + K YRGVR+
Sbjct: 30  WEESFLHQSFDQSFLLSSPTDNYCDDFFAFESSIIKEEGKEATVAA--EEEEKSYRGVRK 87
```

```
Query: 154 RPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP-------- 205
           RPWGKFAAEIRD  + G RVWLGTF+TAE AALAYD+AAF ++GS A+LNFP
Sbjct: 88  RPWGKFAAEIRDSTRKGIRVWLGTFDTAEAAALAYDQAAFALKGSLAVLNFPADVVEESL 147

Query: 206 -----LRVNSGEPDPVRIKSKRS 223
                + +N GE   + +K K S
Sbjct: 148 RKMENVNLNDGESPVIALKRKHS 170


>G1419 (Amino Acid Sequence) (gf=1)
         Length = 221

 Score =  107 bits (266), Expect = 6e-25
 Identities = 50/91 (54%), Positives = 66/91 (71%), Gaps = 5/91 (5%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           +HYRGVR+RPWGK+AAEIRDP K G RVWLGTF+TA +AA  YD+AAF++RGS+A+LNFP
Sbjct: 70  RHYRGVRRRPWGKYAAEIRDPNKKGVRVWLGTFDTAMEAARGYDKAAFKLRGSKAILNFP 129

Query: 206 LRVNSGE-----PDPVRIKSKRSSFSSSNEN 231
           L     E         + +K+KR    + +E+
Sbjct: 130 LEAGKHEDLGDNKKTISLKAKRKRQVTEDES 160


>G1411 (Amino Acid Sequence) (gf=1)
         Length = 248

 Score =  104 bits (260), Expect = 3e-24
 Identities = 56/113 (49%), Positives = 73/113 (64%), Gaps = 10/113 (8%)

Query: 118 ESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGT 177
           +++ +V +     ++ +P         + +HYRGVRQRPWGK+AAEIRDP K  ARVWLGT
Sbjct: 60  DNYPSVYNPQDPNQQVAPTHQDQGDLRRRHYRGVRQRPWGKWAAEIRDP-KKAARVWLGT 118

Query: 178 FETAEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPDPVRIKSKRSSFSSSNE 230
           FETAE AALAYD AA + +GS+A LNFP RV  G           S++ SSN+
Sbjct: 119 FETAESAALAYDEAALKFKGSKAKLNFPERVQLG---------SNSTYYSSNQ 162


>G865 (Amino Acid Sequence) (gf=1)
         Length = 212

 Score =  104 bits (260), Expect = 3e-24
 Identities = 54/84 (64%), Positives = 64/84 (75%), Gaps = 3/84 (3%)

Query: 127 PVKKE--KTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDA 184
           PVK+E  K+ .         + +HYRGVRQRPWGK+AAEIRDP K  ARVWLGTFETAE+A
Sbjct: 16  PVKQELDKSDQHQPDQDQPRRRHYRGVRQRPWGKWAAEIRDP-KKAARVWLGTFETAEEA 74

Query: 185 LAYDRAAFRMRGSRALLNFPLRV 208
           LAYDRAA + +G++A LNFP RV
Sbjct: 75  LAYDRAALKFKGTKAKLNFPERV 98


>G30 (Amino Acid Sequence) (gf=1)
         Length = 133

 Score =  103 bits (257), Expect = 7e-24
 Identities = 50/70 (71%), Positives = 57/70 (81%), Gaps = 1/70 (1%)

Query: 142 AAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRAL 201
           A +GK YRGVR+RPWGK+AAEIRD  K+G RVWLGTF+TAEDAA AYDRAA+ MRG  A+
Sbjct: 15  AEQGK-YRGVRRRPWGKYAAEIRDSRKHGERVWLGTFDTAEDAARAYDRAAYSMRGKAAI 73

Query: 202 LNFPLRVNSG 211
```



```
           LNFP   N G
Sbjct: 74  LNFPHEYNMG 83


>G44 (Amino Acid Sequence) (gf=1)
         Length = 241

 Score =  103 bits (256), Expect = 9e-24
 Identities = 54/113 (47%), Positives = 73/113 (63%), Gaps = 10/113 (8%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           +HYRGVR+RPWGK+AAEIRDP K G R+WLGT++TA +A  AYD+AAF++RG +A+LNFP
Sbjct: 86  RHYRGVRRRPWGKYAAEIRDPNKKGCRIWLGTYDTAVEAGRAYDQAAFQLRGRKAILNFP 145

Query: 206 LRVNSGEPDPVRIKSKRSSFSSSNENGAPKKRRTVAAGGGMDKGLTVKCEVVE 258
           L         VR+ S+  + S     G  K++R   +    +K   VK E  E
Sbjct: 146 L--------DVRVTSE--TCSGEGVIGLGKRKRDKGSPPEEEKAARVKVEEEE 188


>G1795 (Amino Acid Sequence) (gf=1)
         Length = 131

 Score =  100 bits (250), Expect = 5e-23
 Identities = 52/97 (53%), Positives = 67/97 (68%), Gaps = 1/97 (1%)

Query: 140 VTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSR 199
           V A  GK YRGVR+RPWGK+AAEIRD  K+G RVWLGTF+TAE+AA AYD+AA+ MRG
Sbjct: 8   VGAEHGK-YRGVRRRPWGKYAAEIRDSRKHGERVWLGTFDTAEEAARAYDQAAYSMRGQA 66

Query: 200 ALLNFPLRVNSGEPDPVRIKSKRSSFSSSNENGAPKK 236
           A+LNFP   N G           SS +S++ + + ++
Sbjct: 67  AILNFPHEYNMGSGVSSSTAMAGSSSASASASSSSRQ 103


>G1385 (Amino Acid Sequence) (gf=1)
         Length = 336

 Score =  100 bits (249), Expect = 6e-23
 Identities = 60/138 (43%), Positives = 78/138 (56%), Gaps = 10/138 (7%)

Query: 110 PSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKN 169
           PS ++          +P+ +    P SA        K YRGVRQR WGK+ AEIR P +N
Sbjct: 155 PSGRMMMMNMLRQESDLPLTRPPVQPFSAT------KLYRGVRQRHWGKWVAEIRKP-RN 207

Query: 170 GARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP-LRVNSGEPDPVRIKSKRSSFSS- 227
            R+WLGTF+TAE+AA+AYDR AF++RG  A LNFP L +N   EP PV  K   +  +S
Sbjct: 208 RTRLWLGTFDTAEEAAMAYDREAFKLRGETARLNFPELPLNKQEPTPVHQKQCETGTTSE 267

Query: 228 -SNENGAPKKRRTVAAGG 244
            S+  G       +A GG
Sbjct: 268 DSSRRGEDDSSTALAVGG 285


>G1752 (Amino Acid Sequence) (gf=1)
         Length = 243

 Score =  100 bits (248), Expect = 8e-23
 Identities = 50/89 (56%), Positives = 63/89 (70%), Gaps = 7/89 (7%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           + YRGVR+RPWGKFAAEIRD  +NG RVWLGTF+ AE+AALAYD+AAF  +GS A LNFP
Sbjct: 84  RSYRGVRKRPWGKFAAEIRDSTRNGIRVWLGTFDKAEEAALAYDQAAFATKGSLATLNFP 143

Query: 206 LRVNSGEPDPVRIKSKRSSFSSSNENGAP 234
           + V       VR   K+    + ++ G+P
```

```
Sbjct: 144 VEV-------VRESLKKMENVNLHDGGSP 165
```

>G1751 (Amino Acid Sequence) (gf=1)
Length = 268

```
 Score =  100 bits (248), Expect = 8e-23
 Identities = 62/125 (49%), Positives = 72/125 (57%), Gaps = 5/125 (4%)

Query: 124 DSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAED 183
           +S PV K+             K   YRGVRQRPWGKFAAEIRDP K   RVWLGTFETAED
Sbjct: 112 ESSPVAKKAEGGGKIRKRKNKKNGYRGVRQRPWGKFAAEIRDP-KRATRVWLGTFETAED 170

Query: 184 AALAYDRAAFRMRGSRALLNFP-LRVNSGEPDPVR---IKSKRSSFSSSNENGAPKKRRT 239
           AA AYDRAA   RG RA LNFP +   S    PV    I +K S+ +S +   + +  +
Sbjct: 171 AARAYDRAAIGFRGPRAKLNFPPVDYTSSVSSPVAADDIGAKASASASVSATDSVEAEQW 230

Query: 240 VAAGG 244
               GG
Sbjct: 231 NGGGG 235
```

>G2509 (Amino Acid Sequence) (gf=1)
Length = 263

```
 Score = 99.4 bits (246), Expect = 1e-22
 Identities = 48/62 (77%), Positives = 53/62 (85%), Gaps = 1/62 (1%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           +HYRGVRQRPWGK+AAEIRDP K  ARVWLGTFETAE AALAYD AA + +GS+A LNFP
Sbjct: 90  RHYRGVRQRPWGKWAAEIRDPQK-AARVWLGTFETAEAAALAYDNAALKFKGSKAKLNFP 148

Query: 206 LR 207
           R
Sbjct: 149 ER 150
```

>G26 (Amino Acid Sequence) (gf=1)
Length = 218

```
 Score = 99.4 bits (246), Expect = 1e-22
 Identities = 49/69 (71%), Positives = 59/69 (85%), Gaps = 3/69 (4%)

Query: 143 AKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALL 202
           ++ ++YRGVRQRPWGK+AAEIRDP K  ARVWLGTF+TAE+AALAYD+AAF  RG +A L
Sbjct: 65  SRQRNYRGVRQRPWGKWAAEIRDPNK-AARVWLGTFDTAEEAALAYDKAAFEFRGHKAKL 123

Query: 203 NFP--LRVN 209
           NFP  +RVN
Sbjct: 124 NFPEHIRVN 132
```

>G1750 (Amino Acid Sequence) (gf=1)
Length = 335

```
 Score = 98.6 bits (244), Expect = 2e-22
 Identities = 58/137 (42%), Positives = 80/137 (58%), Gaps = 19/137 (13%)

Query: 96  EPSSSSSDEDRSSFPSVKIE----------TPESFAAVDSVPVKKEKTSPVSA-----AV 140
           + + SSSDE+    FP  +++          +  S   V + P K  K   V +      +V
Sbjct: 51  DATDSSSDEEEFLFPRRRVKRLINEIRVEPSSSSTGDVSASPTKDRKRINVDSTVQKPSV 110

Query: 141 TAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRA 200
           +    K YRGVRQRPWGK+AAEIRDP +    R+WLGTF TAE+AA+ YD AA ++RG  A
Sbjct: 111 SGQNQKKYRGVRQRPWGKWAAEIRDPEQR-RRIWLGTFATAEEAAIVYDNAAIKLRGPDA 169
```



```
Query: 201 LLNFPLRVNSGEPDPVR 217
           L NF ++    EP+PV+
Sbjct: 170 LTNFTVQ---PEPEPVQ 183
```

>G864 (Amino Acid Sequence) (gf=1)
Length = 343

```
 Score = 98.6 bits (244), Expect = 2e-22
 Identities = 58/135 (42%), Positives = 79/135 (57%), Gaps = 19/135 (14%)

Query: 128 VKKEKT-SPVSAAVTAAKG-KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAA 185
           VK E T SPV +A T   G K +RGVRQRPWGK+AAEIRDP K   R+WLGT+ TAE+AA
Sbjct: 100 VKSESTVSPVVSATTTTTGEKKFRGVRQRPWGKWAAEIRDPLKR-VRLWLGTYNTAEEAA 158

Query: 186 LAYDRAAFRMRGSRALLNFPLRVNSG------EPDPVRIKSKRSSFSSSNENGAPKKRRT 239
           + YD AA ++RG  AL NF +   +         P PV+ K K+++           K +++
Sbjct: 159 MVYDNAAIQLRGPDALTNFSVTPTTATEKKAPPPSPVKKKKKKNN----------KSKKS 208

Query: 240 VAAGGGMDKGLTVKC 254
           V A   + +  + C
Sbjct: 209 VTASSSISRSSSNDC 223
```

>G4 (Amino Acid Sequence) (gf=1)
Length = 375

```
 Score = 97.8 bits (242), Expect = 4e-22
 Identities = 63/150 (42%), Positives = 84/150 (56%), Gaps = 17/150 (11%)

Query: 91  FHGGWEPSSSSSDEDRSSFPSVK------IETPESFAAVDSVPVKKEKTSPVSAAVTAAK 144
           F G  + S+    ++D   F +VK          P + A V +V     K +  SA
Sbjct: 56  FQGFKDDSAFDCEDDDDVFVNVKPFVFTATTKPVASAFVSTVGSAYAKKTVESAEQAEKS 115

Query: 145 GKH-----YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSR 199
            K      YRG+RQRPWGK+AAEIRDP K G+R WLGTF+TAE+AA AYD AA R+RG++
Sbjct: 116 SKRKRKNQYRGIRQRPWGKWAAEIRDPRK-GSREWLGTFDTAEEAARAYDAAARRIRGTK 174

Query: 200 ALLNFPLRVNSGEPDPVRIKSKRSSFSSSN 229
           A +NFP      E +P  +  KR S  ++N
Sbjct: 175 AKVNFP-----EEKNPSVVSQKRPSAKTNN 199
```

>G2130 (Amino Acid Sequence) (gf=1)
Length = 315

```
 Score = 96.7 bits (239), Expect = 9e-22
 Identities = 57/119 (47%), Positives = 71/119 (58%), Gaps = 14/119 (11%)

Query: 99  SSSSDEDRSSF---PSVKI----------ETPESFAAVDSVPVKKEKTSPVSAAVTAAKG 145
           SSS D+D ++    P VK           E+  S AA      ++E          T+ K
Sbjct: 43  SSSDDDDNNNVTVVPRVKRYVKEIRFCQGESSSSTAARKGKHKEEESVVVEDDVSTSVKP 102

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNF 204
           K YRGVRQRPWGKFAAEIRDP+    R+WLGTF TAE+AA+AYDRAA  ++G +AL NF
Sbjct: 103 KKYRGVRQRPWGKFAAEIRDPSSR-TRIWLGTFVTAEEAAIAYDRAAIHLKGPKALTNF 160
```

>G7 (Amino Acid Sequence) (gf=1)
Length = 192

```
 Score = 96.7 bits (239), Expect = 9e-22
 Identities = 49/78 (62%), Positives = 57/78 (72%), Gaps = 1/78 (1%)
```

43/69 = 62.3%

```
Query: 133 TSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAA 192
           T    ++++  + K YRGVRQRPWGK+AAEIRDP K   RVWLGTFETAE AA AYD AA
Sbjct: 46  TKECESSMSLERPKKYRGVRQRPWGKWAAEIRDPHK-ATRVWLGTFETAEAAARAYDAAA 104

Query: 193 FRMRGSRALLNFPLRVNS 210
           R RGS+A LNFP  V +
Sbjct: 105 LRFRGSKAKLNFPENVGT 122


>G45 (Amino Acid Sequence) (gf=1)
         Length = 349

 Score = 96.3 bits (238), Expect = 1e-21
 Identities = 56/133 (42%), Positives = 73/133 (54%), Gaps = 27/133 (20%)

Query: 107 SSFPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDP 166
           +SFP+     P +   +DS  +              + K + YRGVR+RPWGKFAAEIRD
Sbjct: 124 ASFPNKTNHDPLTNPTIDSCSLSSR---------VSCKTRKYRGVRKRPWGKFAAEIRDS 174

Query: 167 AKNGARVWLGTFETAEDAALAYDRAAFRMRGS-RALLNFPLRV----------------- 208
            +NG RVWLGTF+TAE+AA+AYD+AA R+RG+ +A  NF L
Sbjct: 175 TRNGVRVWLGTFQTAEEAAMAYDKAAVRIRGTQKAHTNFQLETVIKAMEMDCNPNYYRMN 234

Query: 209 NSGEPDPVRIKSK 221
           NS   DP+R   K
Sbjct: 235 NSNTSDPLRSSRK 247


>G19 (Amino Acid Sequence) (gf=1)
         Length = 248

 Score = 96.3 bits (238), Expect = 1e-21
 Identities = 60/141 (42%), Positives = 79/141 (55%), Gaps = 6/141 (4%)

Query: 95  WEPSSSSSDEDRSSFPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQR 154
           W    +S+ +D   F S     P +   V    VKKE+ +          +   YRG+R+R
Sbjct: 28  WSELDASAADDFWGFYSTSKLHPTNQVNVKEEAVKKEQAT--EPGKRRKRKNVYRGIRKR 85

Query: 155 PWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPD 214
           PWGK+AAEIRDP K G RVWLGTF TAE+AA+AYD AA ++RG +A LNFP   +   P
Sbjct: 86  PWGKWAAEIRDPRK-GVRVWLGTFNTAEEAAMAYDVAAKQIRGDKAKLNFP---DLHHPP 141

Query: 215 VRIKSKRSSFSSSNENGAPK 235
           P       SS  S+++  A K
Sbjct: 142 PNYTPPPSSPRSTDQPPAKK 162


>G14 (Amino Acid Sequence) (gf=1)
         Length = 358

 Score = 95.5 bits (236), Expect = 2e-21
 Identities = 62/127 (48%), Positives = 77/127 (59%), Gaps = 13/127 (10%)

Query: 91  FHGGWEPSSSSSDED-------RSSFPSVKIETPE---SFAAVDSVPVKKEKTSPVSAAV 140
           F G  + SS   D+D           P V   TP+   S AA  SV  KK       A
Sbjct: 56  FQGFKDDSSIDCDDDFDVGDVFADVKPFVFTSTPKPAVSAAAEGSVFGKKVTGLDGDAEK 115

Query: 141 TAAKGK--HYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGS 198
           +A + +   YRG+RQRPWGK+AAEIRDP + GAR+WLGTF+TAE+AA AYD AA R+RGS
Sbjct: 116 SANRKRKNQYRGIRQRPWGKWAAEIRDP-REGARIWLGTFKTAEEAARAYDAAARRIRGS 174

Query: 199 RALLNFP 205
           +A +NFP
Sbjct: 175 KAKVNFP 181
```

42/70 = 60%


```
>G1008 (Amino Acid Sequence) (gf=1)
         Length = 294

 Score = 94.4 bits (233), Expect = 4e-21
 Identities = 55/117 (47%), Positives = 66/117 (56%), Gaps = 9/117 (7%)

Query: 96  EPSSSSSDEDRSSFPSVKIETPE-SFAAVDSVPVKKEKTSPVSAAVTAAKGKH------- 147
           + SS   ++++S  P VK    E  F   D  P    K    S A  A  G
Sbjct: 39  DSSSDEENDNKSVAPRVKRYVDEIRFCDEDDEPKPARKAKKKSPAAAAENGGDLVKSVVK 98

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNF 204
           YRGVRQRPWGKFAAEIRDP+    R+WLGTF TAE+AA+ YDRAA R++G  A  NF
Sbjct: 99  YRGVRQRPWGKFAAEIRDPSSR-TRLWLGTFATAEEAAIGYDRAAIRIKGHNAQTNF 154


>G1421 (Amino Acid Sequence) (gf=1)
         Length = 287

 Score = 94.0 bits (232), Expect = 6e-21
 Identities = 56/121 (46%), Positives = 76/121 (62%), Gaps = 16/121 (13%)

Query: 99  SSSSDEDRSSFPSV-----KIETPESFAAVDSV------PVKKEKTS---PVSAAVTAAK 144
           SSS +E+   F ++      +++       +DSV      P+KK++       V   VT A
Sbjct: 25  SSSDEEEEVDFDALSTKRRRVKKYVKEVVLDSVVSDKEKPMKKKRKKRVVTVPVVVTTAT 84

Query: 145 GKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNF 204
            K +RGVRQRPWGK+AAEIRDP++   RVWLGTF+TAE+AA+ YD AA ++RG  A LNF
Sbjct: 85  RK-FRGVRQRPWGKWAAEIRDPSRR-VRVWLGTFDTAEEAAIVYDNAAIQLRGPNAELNF 142

Query: 205 P 205
           P
Sbjct: 143 P 143


>G1794 (Amino Acid Sequence) (gf=1)
         Length = 391

 Score = 93.6 bits (231), Expect = 7e-21
 Identities = 51/89 (57%), Positives = 59/89 (65%), Gaps = 6/89 (6%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           + YRGVRQRPWGK+AAEIRDP K  ARVWLGTF+ AE AA AYD AA R RG++A LNFP
Sbjct: 183 RRYRGVRQRPWGKWAAEIRDPFK-AARVWLGTFDNAESAARAYDEAALRFRGNKAKLNFP 241

Query: 206 -----LRVNSGEPDPVRIKSKRSSFSSSN 229
                +R  S E  PV   + +    S N
Sbjct: 242 ENVKLVRPASTEAQPVHQTAAQRPTQSRN 270


>G440 (Amino Acid Sequence) (gf=1)
         Length = 354

 Score = 92.8 bits (229), Expect = 1e-20
 Identities = 66/177 (37%), Positives = 88/177 (49%), Gaps = 31/177 (17%)

Query: 56  SFSKLYPCFTESWGDLPLKENDSEDML-----VYGILNDAFHGGWEPSS-------SSSD 103
           S  K+  C+T+          + D ED L      V    +N+       EPS        S  D
Sbjct: 35  SLKKVSICYTDPDATDSSSDEDEEDFLFPRRRVKRFVNEITV---EPSCNNVVTGVSMKD 91

Query: 104 EDRSSFPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEI 163
           + R S  S + ++P S              +  P +     + + K +RGVRQRPWGK+AAEI
Sbjct: 92  RKRLSSSSDETQSPAS-----------SRQRPNNKVSVSGQIKKFRGVRQRPWGKWAAEI 140

Query: 164 RDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPL----RVNSGEPDPV 216
           RDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPL    RVNSGEPDPV
```

```
          RDP +   R+WLGTFETAE+AA+ YD AA R+RG  AL NF +         EP+PV
Sbjct: 141 RDPEQR-RRIWLGTFETAEEAAVVYDNAAIRLRGPDALTNFSIPPQEEEEEEEPEPV 196


>G31 (Amino Acid Sequence) (gf=1)
         Length = 200

 Score = 92.0 bits (227), Expect = 2e-20
 Identities = 46/82 (56%), Positives = 55/82 (66%), Gaps = 1/82 (1%)

Query: 125 SVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDA 184
           S+ V +          T  K  H+RGVR+RPWG++AAEIRDP K   RVWLGTF+TAE+A
Sbjct: 10  SIAVSEGGGGKTMTMTTMRKEVHFRGVRKRPWGRYAAEIRDPGKK-TRVWLGTFDTAEEA 68

Query: 185 ALAYDRAAFRMRGSRALLNFPL 206
           A AYD AA   RGS+A  NFPL
Sbjct: 69  ARAYDTAAREFRGSKAKTNFPL 90


>G25 (Amino Acid Sequence) (gf=1)
         Length = 171

 Score = 92.0 bits (227), Expect = 2e-20
 Identities = 51/94 (54%), Positives = 64/94 (67%), Gaps = 11/94 (11%)

Query: 117 PESFAAVDSVPVKKEKTSPVSAAVTAAKGKH-----YRGVRQRPWGKFAAEIRDPAKNGA 171
           P    +V S    ++K  PVS +     GK      YRG+RQRPWGK+AAEIRDP+K G
Sbjct: 19  PSQLGSVSS----RKKRKPVSVS-EERDGKRERKNLYRGIRQRPWGKWAAEIRDPSK-GV 72

Query: 172 RVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           RVWLGTF+TA++AA AYD AA ++RG +A LNFP
Sbjct: 73  RVWLGTFKTADEAARAYDVAAIKIRGRKAKLNFP 106


>G1763 (Amino Acid Sequence) (gf=1)
         Length = 314

 Score = 91.3 bits (225), Expect = 4e-20
 Identities = 53/113 (46%), Positives = 68/113 (59%), Gaps = 11/113 (9%)

Query: 137 SAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMR 196
           S+   +   K YRGVRQR WGK+ AEIR P +N  R+WLGTF+TAE+AALAYDRAAF++R
Sbjct: 132 SSGSVSKPAKLYRGVRQRHWGKWVAEIRLP-RNRTRLWLGTFDTAEEAALAYDRAAFKLR 190

Query: 197 GSRALLNFP-LRVNSGE---------PDPVRIKSKRSSPSSSNENGAPKKRRT 239
           G  A LNFP LR  +G          P    + +K  +  +  +N   K  RT
Sbjct: 191 GDSARLNFPALRYQTGSSPSDVGEYGPIQAAVDAKLEAILAEPKNQPGKTERT 243


>G1384 (Amino Acid Sequence) (gf=1)
         Length = 314

 Score = 91.3 bits (225), Expect = 4e-20
 Identities = 53/113 (46%), Positives = 68/113 (59%), Gaps = 11/113 (9%)

Query: 137 SAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMR 196
           S+   +   K YRGVRQR WGK+ AEIR P +N  R+WLGTF+TAE+AALAYDRAAF++R
Sbjct: 132 SSGSVSKPAKLYRGVRQRHWGKWVAEIRLP-RNRTRLWLGTFDTAEEAALAYDRAAFKLR 190

Query: 197 GSRALLNFP-LRVNSGE---------PDPVRIKSKRSSPSSSNENGAPKKRRT 239
           G  A LNFP LR  +G          P    + +K  +  +  +N   K  RT
Sbjct: 191 GDSARLNFPALRYQTGSSPSDTGEYGPIQAAVDAKLEAILAEPKNQPGKTERT 243


>G974 (Amino Acid Sequence) (gf=1)
```



```
         Length = 261

 Score = 91.3 bits (225), Expect = 4e-20
 Identities = 61/172 (35%), Positives = 92/172 (53%), Gaps = 16/172 (9%)

Query: 97  PSSSSSDEDRSSFPSVKIETPESFAA--------VDSVPVKKEKTSPVSAAVTAAKGKHY 148
           P   S   D S+F +  +  P S+ +         +   PV  ++T   +A  T   K Y
Sbjct: 29  PFIKSPSNDSSAP-AFSLPAPISYGSDLHSPSHHLSPKPVSMKQTGTSAAKPT----KLY 83

Query: 149 RGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLRV 208
           RGVRQR WGK+ AEIR P +N  R+WLGTF+TAE+AALAYD+AA+++RG  A LNFP
Sbjct: 84  RGVRQRHWGKWVAEIRLP-RNRTRLWLGTFDTAEEAALAYDKAAYKLRGDFARLNFPDLR 142

Query: 209 NSGEPDPVR--IKSKRSSPSSSNENGAPKKRRTVAAGGGMDKGLTVKCEVVE 258
           ++ E  P++  + +K  +   +      K+ R+        +  TV  ++ E
Sbjct: 143 HNDEYQPLQSSVDAKLEAICQNLAETTQKQVRSTKKSSSRKRSSTVAVKLPE 194


>G33 (Amino Acid Sequence) (gf=1)
         Length = 245

 Score = 90.9 bits (224), Expect = 5e-20
 Identities = 50/93 (53%), Positives = 58/93 (61%), Gaps = 5/93 (5%)

Query: 118 ESFAAVDSVPVKKEKTSPVSAAVT----AAKGKHYRGVRQRPWGKFAAEIRDPAKNGARV 173
           E    V    VKK K    + AV      +K   YRGVR+RPWG++AAEIRDP K   RV
Sbjct: 19  EKSKEVSDKGVKKRKNVTKALAVNDGGEKSKEVRYRGVRRRPWGRYAAEIRDPVKK-KRV 77

Query: 174 WLGTFETAEDAALAYDRAAFRMRGSRALLNFPL 206
           WLGTF T E+AA AYD AA R RGS+A  NFPL
Sbjct: 78  WLGSFNTGEEAARAYDSAAIRFRGSKATTNFPL 110


>G1932 (Amino Acid Sequence) (gf=1)
         Length = 189

 Score = 90.5 bits (223), Expect = 6e-20
 Identities = 44/68 (64%), Positives = 53/68 (77%), Gaps = 1/68 (1%)

Query: 138 AAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRG 197
           A+ T A+  HYRGVR+RPWG++AAEIRDP K   RVWLGTP+T E+AALAYD AA  +RG
Sbjct: 2   ASTTCAREVHYRGVRKRPWGRYAAEIRDPWKK-TRVWLGTFDTPEEAALAYDGAARFLRG 60

Query: 198 SRALLNFP 205
           +A  NFP
Sbjct: 61  IKAKTNFP 68


>G1005 (Amino Acid Sequence) (gf=1)
         Length = 225

 Score = 90.5 bits (223), Expect = 6e-20
 Identities = 47/88 (53%), Positives = 61/88 (68%), Gaps = 1/88 (1%)

Query: 136 VSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRM 195
           V+A   + K   +RGVR+RPWG+FAAEIRDP K  ARVWLGTF++AE+AA AYD AA  +
Sbjct: 16  VAAINGSVKEIRFRGVRKRPWGRFAAEIRDPWKK-ARVWLGTFDSAEEAARAYDSAARNL 74

Query: 196 RGSRALLNFPLRVNSGEPDPVRIKSKRS 223
           RG +A  NFP+  +S  P  +R    R+
Sbjct: 75  RGPKAKTNFPIDSSSPPPPNLRFNQIRN 102


>G439 (Amino Acid Sequence) (gf=1)
         Length = 279
```

```
 Score = 89.4 bits (220), Expect = 1e-19
 Identities = 56/122 (45%), Positives = 70/122 (56%), Gaps = 7/122 (5%)

Query: 133 TSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAA 192
           TS     V     K YRGVRQR WGK+ AEIR P KN  R+WLGTFETA++AALAYD+AA
Sbjct: 98  TSMKKIDVATKPVKLYRGVRQRQWGKWVAEIRLP-KNRTRLWLGTFETAQEAALAYDQAA 156

Query: 193 FRMRGSRALLNFPLRVNSGEPDPV---RIKSKRSSPSSSNENGAP---KKRRTVAAGGGM 246
            ++RG  A LNFP  V  G    +    I +K  S  +S++   P   K+ +T    G
Sbjct: 157 HKIRGDNARLNFPDIVRQGHYKQILSPSINAKIESICNSSDLPLPQIEKQNKTEEVLSGF 216

Query: 247 DK 248
           K
Sbjct: 217 SK 218


>G2135 (Amino Acid Sequence) (gf=1)
         Length = 262

 Score = 88.2 bits (217), Expect = 3e-19
 Identities = 50/103 (48%), Positives = 65/103 (62%), Gaps = 13/103 (12%)

Query: 103 DEDRSSPPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAE 162
           DE  S  P V   T +  A+  +   KK+++S             Y+G+R+RPWG++AAE
Sbjct: 43  DELESDEPPVFSSTHKHHASGSASDGKKKQSS------------RYKGIRRRPWGRWAAE 90

Query: 163 IRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           IRDP K G RVWLGTF TAE+AA AYD  A R+RG++A LNFP
Sbjct: 91  IRDPIK-GVRVWLGTFNTAEEAARAYDLEAKRIRGAKAKLNFP 132


>G1380 (Amino Acid Sequence) (gf=1)
         Length = 244

 Score = 88.2 bits (217), Expect = 3e-19
 Identities = 45/79 (56%), Positives = 56/79 (69%), Gaps = 2/79 (2%)

Query: 135 PVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFR 194
           PV+A   + K   YRGVR+RPWG+FAAEIRDP K  +RVWLGTF++A DAA AYD AA
Sbjct: 15  PVTAG-GSVKEPRYRGVRKRPWGRFAAEIRDPLKK-SRVWLGTFDSAVDAARAYDTAARN 72

Query: 195 MRGSRALLNFPLRVNSGEP 213
           +RG +A  NFP+  +   P
Sbjct: 73  LRGPKAKTNFPIDCSPSSP 91


>G625 (Amino Acid Sequence) (gf=1)
         Length = 328

 Score = 88.2 bits (217), Expect = 3e-19
 Identities = 63/130 (48%), Positives = 71/130 (54%), Gaps = 16/130 (12%)

Query: 134 SPVSAAVTAAKGK--------HYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAA 185
           S  S+A     KGK             YRGVRQR WGK+ AEIR+P K   R WLGTF TAEDAA
Sbjct: 33  SSTSSAQRKRKGKGGPDNSKFRYRGVRQRSWGKWVAEIREPRKR-TRKWLGTFATAEDAA 91

Query: 186 LAYDRAAFRMRGSRALLNF----PLRVNSGEPDPVRIKSKRSSPSSSNENGAPKKRRTVA 241
            AYDRAA  + GSRA LN      P  V+S     V   S  S+ SSS +   P   R  A
Sbjct: 92  RAYDRAAVYLYGSRAQLNLTPSSPSSVSSSS-SSVSAASSPSTSSSSTQTLRPLLPRPAA 150

Query: 242 A--GGGMDKG 249
           A  GGG + G
Sbjct: 151 ATVGGGANFG 160
```



```
>G1748 (Amino Acid Sequence) (gf=1)
         Length = 261

 Score = 87.4 bits (215), Expect = 5e-19
 Identities = 47/82 (57%), Positives = 57/82 (69%), Gaps = 2/82 (2%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLR 207
           YRGVRQR WGK+ AEIR P KN  R+WLGTFETAE AALAYD+AAF++RG  A LNFP
Sbjct: 71  YRGVRQRHWGKWVAEIRLP-KNRTRLWLGTFETAEKAALAYDQAAFQLRGDIAKLNFPNL 129

Query: 208 VNSG-EPDPVRIKSKRSSPSSS 228
           V SG  P P  + +K  +  S
Sbjct: 130 IHEDMNPLPSSVDTKLQAICKS 151


>G1845 (Amino Acid Sequence) (gf=1)
         Length = 292

 Score = 87.0 bits (214), Expect = 7e-19
 Identities = 41/60 (68%), Positives = 50/60 (83%), Gaps = 1/60 (1%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           K YRGVRQR WGK+ AEIR P ++ AR+WLGTF+TAE+AA+AYDR AF++RG  A LNFP
Sbjct: 141 KLYRGVRQRQWGKWVAEIRKP-RSRARLWLGTFDTAEEAAMAYDRQAFKLRGHSATLNFP 199


>G2059 (Amino Acid Sequence) (gf=1)
         Length = 343

 Score = 86.3 bits (212), Expect = 1e-18
 Identities = 54/111 (48%), Positives = 66/111 (58%), Gaps = 9/111 (8%)

Query: 130 KEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYD 189
           K KT  V    T  K   YRGVRQR WGK+ AEIR P +N  RVWLGTFETAE AA+AYD
Sbjct: 171 KSKTRKVVQTTTPTK--LYRGVRQRHWGKWVAEIRLP-RNRTRVWLGTFETAEQAAMAYD 227

Query: 190 RAAFRMRGSRALLNFPLRVNSGEPDPVR------IKSKRSSPSSSNENGAP 234
           RAA+ +RG  A LNFP    +   +R       ++SK    SSS  + +P
Sbjct: 228 TAAYILRGEPAHLNFPDLKHQLKSGSLRCMIASLLESKIQQISSSQVSNSP 278


>G1020 (Amino Acid Sequence) (gf=1)
         Length = 185

 Score = 86.3 bits (212), Expect = 1e-18
 Identities = 45/78 (57%), Positives = 53/78 (67%), Gaps = 1/78 (1%)

Query: 133 TSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAA 192
           T+P      AAK   YRGVR+RPWG++AAEIRDP K   RVWLGTF+TA+ AA AYD AA
Sbjct: 16  TNPTHHESNAAKEIRYRGVRKRPWGRYAAEIRDPVKK-TRVWLGTFDTAQQAARAYDAAA 74

Query: 193 FRMRGSRALLNFPLRVNS 210
              RG +A  NF + V S
Sbjct: 75  RDFRGVKAKTNFGVIVGS 92


>G6 (Amino Acid Sequence) (gf=1)
         Length = 222

 Score = 86.3 bits (212), Expect = 1e-18
 Identities = 42/63 (66%), Positives = 49/63 (77%), Gaps = 1/63 (1%)

Query: 143 AKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALL 202
           AK   YRGVR+RPWG++AAEIRDP K   RVWLGTF+TAE+AA AYD AA   RG++A
```

```
Sbjct: 20  AKEIRYRGVRKRPWGRYAAEIRDPGKK-TRVWLGTFDTAEEAARAYDTAARDPRGAKAKT 78

Query: 203 NFP 205
           NFP
Sbjct: 79  NFP 81
```

>G1796 (Amino Acid Sequence) (gf=1)
Length = 328

```
 Score = 85.9 bits (211), Expect = 2e-18
 Identities = 52/128 (40%), Positives = 66/128 (50%), Gaps = 9/128 (7%)

Query: 110 PSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKN 169
           P  K  T  + + V         K + V+ A    +     YRGVR+RPWG++AAEIRDP
Sbjct: 19  PLTKFPTEPTASPVSRNRKLSSKDTTVTIAGAGSSTTRYRGVRRRPWGRYAAEIRDPMSK 78

Query: 170 GARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPDDPVRIKSKRSSFSSSN 229
             R WLGTF+TAE AA AYD AA    RG++A  NF          P+P          FS SN
Sbjct: 79  ERR-WLGTFDTAEQAACAYDSAARAFRGAKARTNFTYPTAVIMPEP--------RFSFSN 129

Query: 230 ENGAPKKR 237
           +  +P  R
Sbjct: 130 KKSSPSAR 137
```

>G5 (Amino Acid Sequence) (gf=1)
Length = 334

```
 Score = 85.9 bits (211), Expect = 2e-18
 Identities = 42/68 (61%), Positives = 53/68 (77%), Gaps = 1/68 (1%)

Query: 138 AAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRG 197
           + V +    K YRGVRQR WGK+ AEIR P +N  R+WLGTF+TAE+AALAYD+AA+++RG
Sbjct: 142 SGVPSKPTKLYRGVRQRHWGKWVAEIRLP-RNRTRLWLGTFDTAEEAALAYDKAAYKLRG 200

Query: 198 SRALLNFP 205
             A LNFP
Sbjct: 201 DFARLNFP 208
```

>G2113 (Amino Acid Sequence) (gf=1)
Length = 166

```
 Score = 85.1 bits (209), Expect = 3e-18
 Identities = 39/61 (63%), Positives = 48/61 (77%), Gaps = 1/61 (1%)

Query: 145 GKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNF 204
           G  YRGVR+RPWG++AAEIRDP K   +RVWLGTF+T E+AA AYD+ A    RG++A  NF
Sbjct: 17  GVRYRGVRKRPWGRYAAEIRDPFKK-SRVWLGTFDTPEEAARAYDKRAIEFRGAKAKTNF 75

Query: 205 P 205
           P
Sbjct: 76  P 76
```

>G2571 (Amino Acid Sequence) (gf=1)
Length = 335

```
 Score = 84.7 bits (208), Expect = 3e-18
 Identities = 43/70 (61%), Positives = 52/70 (73%), Gaps = 1/70 (1%)

Query: 136 VSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRM 195
           ++A     K K YRGVRQR WGK+ AEIR P +N  RVWLGT++TAE AAAYAYDRAAYKL 182
Sbjct: 124 LAAGRVTKKKKLYRGVRQRHWGKWVAEIRLP-QNRMRVWLGTYDTAEAAAYAYDRAAYKL
```



```
Query: 196 RGSRALLNFP 205
           RG  A LNFP
Sbjct: 183 RGEYARLNFP 192
```

>G36 (Amino Acid Sequence) (gf=1)
Length = 272

```
 Score = 84.3 bits (207), Expect = 4e-18
 Identities = 41/60 (68%), Positives = 49/60 (81%), Gaps = 1/60 (1%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           K YRGVRQR WGK+ AEIR P KN  R+WLGTF+TAE+AA+AYD AA+++RG  A LNFP
Sbjct: 91  KLYRGVRQRHWGKWVAEIRLP-KNRTRLWLGTFDTAEEAAMAYDLAAYKLRGEFARLNFP 149
```

>G32 (Amino Acid Sequence) (gf=1)
Length = 211

```
 Score = 82.4 bits (202), Expect = 2e-17
 Identities = 40/60 (66%), Positives = 47/60 (77%), Gaps = 1/60 (1%)

Query: 145 GKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNF 204
           G  + GVR+RPWG++AAEIRDP     R WLGTF+TAE+AALAYDRAA  MRG+RA  NF
Sbjct: 17  GTRFLGVRRRPWGRYAAEIRDPTTK-ERHWLGTFDTAEEAALAYDRAARSMRGTRARTNF 75
```

>G35 (Amino Acid Sequence) (gf=1)
Length = 159

```
 Score = 82.0 bits (201), Expect = 2e-17
 Identities = 57/127 (44%), Positives = 68/127 (52%), Gaps = 16/127 (12%)

Query: 98  SSSSSDEDRSSFPSVKIETPESFAAVDSVPVKKEKT--------SPVSAAVTAAKGKHYR 149
           SSSS DE   S      +  +      S P +  KT        S ++A  T A  K +R
Sbjct: 17  SSSSEDESPPSSRRRGKKLVKEIVIDHSDPPEVGKTRFKIRIPASLLAARNTTANKKKFR 76

Query: 150 GVRQRPWGKFAAEIRDPAKNGA--RVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLR 207
           GVRQRPWGK+AAEIR     G   R+WLGTFETAE+AALAYD AA ++ G  A  NF
Sbjct: 77  GVRQRPWGKWAAEIRCGRVKGRPERIWLGTFETAEEAALAYDNAAIQLIGPDAPTNF--- 133

Query: 208 VNSGEPD 214
              G PD
Sbjct: 134 ---GRPD 137
```

>G1415 (Amino Acid Sequence) (gf=1)
Length = 206

```
 Score = 81.3 bits (199), Expect = 4e-17
 Identities = 49/129 (37%), Positives = 72/129 (54%), Gaps = 14/129 (10%)

Query: 129 KKEKTSPVSAAVTAAKGKH--------YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFET 180
           KK++T   S+     +GK          Y+GVRQR WGK+ AEIR+P + GAR+WLGTF+T
Sbjct: 15  KKQRTVQASSRKGCMRGKGGPDNASCTYKGVRQRTWGKWVAEIREPNR-GARLWLGTFDT 73

Query: 181 AEDAALAYDRAAFRMRGSRALLNFPLRVNS-----GEPDPVRIKSKRSSFSSSNENGAPK 235
           + +AALAYD AA ++ G  A LN P  + S       P      S  +   SS+++ +P
Sbjct: 74  SREAALAYDSAARKLYGPEAHLNLPESLRSYPKTASSPASQTTPSSNTGGKSSSDSESPC 133

Query: 236 KRRTVAAGG 244
               +++ G
Sbjct: 134 SSNEMSSCG 142
```

```
>G23 (Amino Acid Sequence) (gf=1)
         Length = 236

 Score = 80.1 bits (196), Expect = 8e-17
 Identities = 44/109 (40%), Positives = 68/109 (62%), Gaps = 1/109 (0%)

Query: 97  PSSSSSDEDRSSFPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPW 156
           P+SSSSD+  ++  +    + E     ++   K+ + S  +++ ++ +   YRGVR R W
Sbjct: 10  PTSSSSDQSDATTTTSTHLSEEEAPPRNNNTRKRRRDSSSASSSSSMQHPVYRGVRMRSW 69

Query: 157 GKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           GK+ +EIR P K   R+WLGTF TA+ AA A+D AA  ++GS A+LNFP
Sbjct: 70  GKWVSEIRQPRKK-TRIWLGTFVTADMAARAHDVAALTIKGSSAVLNFP 117


>G12 (Amino Acid Sequence) (gf=1)
         Length = 196

 Score = 80.1 bits (196), Expect = 8e-17
 Identities = 43/92 (46%), Positives = 55/92 (59%), Gaps = 6/92 (6%)

Query: 114 IETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARV 173
           +ET    A V S P        V+A  T  + K Y+G+R R WGK+ AEIR+P K  +R+
Sbjct: 1   METATEVATVVSTPAVT-----VAAVATRKRDKPYKGIRMRKWGKWVAEIREPNKR-SRI 54

Query: 174 WLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           WLG++ T E AA AYD A F +RG  A LNFP
Sbjct: 55  WLGSYSTPEAAARAYDTAVFYLRGPSARLNFP 86


>G1141 (Amino Acid Sequence) (gf=1)
         Length = 330

 Score = 79.7 bits (195), Expect = 1e-16
 Identities = 42/63 (66%), Positives = 46/63 (72%), Gaps = 1/63 (1%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLR 207
           +RGVRQR WGK+ AEIR+P K G R+WLGTF TAE AA AYD AA  M GS A LNFP
Sbjct: 78  FRGVRQRIWGKWVAEIREP-KIGTRLWLGTFPTAEKAASAYDEAATAMYGSLARLNFPQS 136

Query: 208 VNS 210
           V S
Sbjct: 137 VGS 139


>G2138 (Amino Acid Sequence) (gf=1)
         Length = 161

 Score = 79.3 bits (194), Expect = 1e-16
 Identities = 54/141 (38%), Positives = 68/141 (47%), Gaps = 22/141 (15%)

Query: 96  EPSSSSSDEDRSS--------------FPSVKIETPESFAAVDSVPVKKEKTSPVSAAVT 141
           E + SSSDED                    + I+  +S   +D    + +   P     T
Sbjct: 13  EATDSSSDEDTEERGGASQTRRRGKRLVKEIVIDPSDSADKLDVCKTRFKIRIPAEFLKT 72

Query: 142 AAKGKHYRGVRQRPWGKFAAEIRDPAKNGA------RVWLGTFETAEDAALAYDRAAFRM 195
           A   K YRGVRQRPWGK+ AEIR     GA      R+WLGTF TAE+AALAYD A+ ++
Sbjct: 73  AKTEKKYRGVRQRPWGKWVAEIR--CGRGACKGRRDRLWLGTFNTAEEAALAYDNASIKL 130

Query: 196 RGSRALLNFPLRVNSGEPDPV 216
            G  A  NF L   + E   V
Sbjct: 131 IGPHAPTNFGLPAENQEDKTV 151
```



```
>G38 (Amino Acid Sequence) (gf=1)
         Length = 335

 Score = 79.3 bits (194), Expect = 1e-16
 Identities = 45/91 (49%), Positives = 55/91 (59%), Gaps = 9/91 (9%)

Query: 123 VDSVPVKKEKTSPVSAAVTAAKGK--------HYRGVRQRPWGKFAAEIRDPAKNGARVW 174
           V+ V  KK K      +      KGK          +RGVRQR WGK+ AEIR+P + G+R+W
Sbjct: 46  VEEVSTKKRKVPAKGSKKGCMKGKGGPENSRCSFRGVRQRIWGKWVAEIREPNR-GSRLW 104

Query: 175 LGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           LGTF TA++AA AYD AA  M G  A LNFP
Sbjct: 105 LGTFPTAQEAASAYDEAAKAMYGPLARLNFP 135


>G975 (Amino Acid Sequence) (gf=1)
         Length = 199

 Score = 79.0 bits (193), Expect = 2e-16
 Identities = 45/89 (50%), Positives = 54/89 (60%), Gaps = 2/89 (2%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           K +RGVRQR WG + AEIR P      R+WLGTFETAE+AA AYD AA  M G  A  NFP
Sbjct: 5   KKFRGVRQRHWGSWVAEIRHPLLK-RRIWLGTFETAEEAARAYDEAAVLMSGRNAKTNFP 63

Query: 206 L-RVNSGEPDPVRIKSKRSSFSSSNENGA 233
           L   N+GE    +     SS  SS+ + +
Sbjct: 64  LNNNNTGETSEGKTDISASSTMSSSTSSS 92


>G1800 (Amino Acid Sequence) (gf=1)
         Length = 277

 Score = 78.2 bits (191), Expect = 3e-16
 Identities = 37/58 (63%), Positives = 45/58 (76%), Gaps = 1/58 (1%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           YRGVRQR WGK+ AEIR+P K  AR+WLG+F TAE+AA+AYD AA ++ G  A LN P
Sbjct: 28  YRGVRQRTWGKWVAEIREPKKR-ARLWLGSFATAEEAAMAYDEAALKLYGHDAYLNLP 84


>G442 (Amino Acid Sequence) (gf=1)
         Length = 244

 Score = 76.3 bits (186), Expect = 1e-15
 Identities = 42/92 (45%), Positives = 53/92 (56%), Gaps = 7/92 (7%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKN-------GARVWLGTFETAEDAALAYDRAAFRMRGSRA 200
           +RGVRQR WGK+ AEIR+P  +         R+WLGTF TA +AALAYDRAA  M G  A
Sbjct: 70  FRGVRQRVWGKWVAEIREPVSHRGANSSRSKRLWLGTFATAAEAALAYDRAASVMYGPYA 129

Query: 201 LLNFPLRVNSGEPDPVRIKSKRSSFSSSNENG 232
            LNFP  + G       +S    +N+ G
Sbjct: 130 RLNFPEDLGGGRKKDEEAESSGGYWLETNKAG 161


>G1754 (Amino Acid Sequence) (gf=1)
         Length = 341

 Score = 75.9 bits (185), Expect = 2e-15
 Identities = 49/126 (38%), Positives = 62/126 (48%), Gaps = 9/126 (7%)

Query: 127 PVKKEKTSPVSAAVTAAKGK--------HYRGVRQRPWGKFAAEIRDPAKNGARVWLGTF 178
           P    K  P +     KGK          YRGVRQR WGK+ AEIR+P   GAR+WLGTF
Sbjct: 43  PKSIRKPPPKGSRKGCMKGKGGPENGICDYRGVRQRRWGKWVAEIREP-DGGARLWLGTF 101
```

```
Query: 179 ETAEDAALAYDRAAFRMRGSRALLNFPLRVNSGEPDPVRIKSKRSSFSSSNENGAPKKRR 238
           ++ +AALAYD AA  + G  A LN P   N              S  + S+E+    +
Sbjct: 102 SSSYEAALAYDEAAKAIYGQSARLNLPEITNRSSSTAATATVSGSVTAFSDESEVCARED 161

Query: 239 TVAAGG 244
           T A+ G
Sbjct: 162 TNASSG 167


>G1090 (Amino Acid Sequence) (gf=1)
         Length = 184

 Score = 75.9 bits (185), Expect = 2e-15
 Identities = 36/63 (57%), Positives = 46/63 (72%), Gaps = 1/63 (1%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLR 207
           YRG+R+R WGK+ +EIR+P K    R+WLG++ETAE AA AYD AA  +RG    LNFP
Sbjct: 20  YRGIRRRKWGKWVSEIREPGKK-TRIWLGSYETAEMAAAAYDAAALHLRGRGTNLNFPEL 78

Query: 208 VNS 210
           V+S
Sbjct: 79  VDS 81


>G1007 (Amino Acid Sequence) (gf=1)
         Length = 225

 Score = 75.9 bits (185), Expect = 2e-15
 Identities = 43/83 (51%), Positives = 53/83 (63%), Gaps = 2/83 (2%)

Query: 123 VDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAE 182
           VDS     E +S      T  KG  YRG R R WGK+ +EIR+P K  +R+WLGTF TAE
Sbjct: 2   VDSHGSDTECSSKKKKEKTKEKGV-YRGARMRSWGKWVSEIREPRKK-SRIWLGTFPTAE 59

Query: 183 DAALAYDRAAFRMRGSRALLNFP 205
            AA A+D AA  ++GS A+LNFP
Sbjct: 60  MAARAHDVAALSIKGSSAILNFP 82


>G2299 (Amino Acid Sequence) (gf=1)
         Length = 236

 Score = 75.5 bits (184), Expect = 2e-15
 Identities = 41/79 (51%), Positives = 54/79 (67%), Gaps = 4/79 (5%)

Query: 145 GKH--YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALL 202
           GKH  YRGVR R WGK+ +EIR+P K  +R+WLGTF T E AA A+D AA  ++G+ A+L
Sbjct: 46  GKHPVYRGVRMRNWGKWVSEIREPRKK-SRIWLGTFPTPEMAARAHDVAALSIKGTAAIL 104

Query: 203 NFPLRVNSGEPDPVRIKSK 221
           NFP   +S  P PV +  +
Sbjct: 105 NFPELADS-FPRPVSLSPR 122


>G16 (Amino Acid Sequence) (gf=1)
         Length = 218

 Score = 75.5 bits (184), Expect = 2e-15
 Identities = 48/116 (41%), Positives = 71/116 (60%), Gaps = 21/116 (18%)

Query: 129 KKEKTSPVSAAVTAAKGKH--YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAAL 186
           ++EK  PV  +     GKH  YRGVR+R WGK+ +EIR+P K  +R+WLGTF + E AA
Sbjct: 20  EEEKKKPVKDS-----GKHPVYRGVRKRNWGKWVSEIREPRKK-SRIWLGTFPSPEMAAR 73
```



```
Query: 187 AYDRAAFRMRGSRALLNFP------LRVNSGEPDPVRIKSKRS-------SFSSSN 229
           A+D AA  ++G+ A+LNFP        R +S  P  +++ + ++         SFSSS+
Sbjct: 74  AHDVAALSIKGASAILNFPDLAGSFPRPSSLSPRDIQVAALKAAHMETSQSFSSSS 129


>G2134 (Amino Acid Sequence) (gf=1)
         Length = 202

 Score = 75.1 bits (183), Expect = 3e-15
 Identities = 36/58 (62%), Positives = 42/58 (72%), Gaps = 1/58 (1%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           YRGVR+R WGK+ +EIR+P      R+WLG+FET E AA AYD AAF  RG   A LNFP
Sbjct: 24  YRGVRKRKWGKWVSEIREPGTKN-RIWLGSFETPEMAATAYDVAAFHFRGREARLNFP 80


>G977 (Amino Acid Sequence) (gf=1)
         Length = 181

 Score = 75.1 bits (183), Expect = 3e-15
 Identities = 38/65 (58%), Positives = 44/65 (67%), Gaps = 1/65 (1%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           + +RGVRQR WG + +EIR P      R+WLGTFETAEDAA AYD AA  M G RA  NFP
Sbjct: 6   QRFRGVRQRHWGSWVSEIRHPLLK-TRIWLGTFETAEDAARAYDEAARLMCGPRARTNFP 64

Query: 206 LRVNS 210
             V+N+
Sbjct: 65  YNPNA 69


>G24 (Amino Acid Sequence) (gf=1)
         Length = 176

 Score = 74.7 bits (182), Expect = 4e-15
 Identities = 41/84 (48%), Positives = 52/84 (61%), Gaps = 10/84 (11%)

Query: 131 EKTSPVSAAVTAAKG---------KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETA 181
           E  + V+A VTAA          K Y+G+R R WGK+ AEIR+P K  +R+WLG+FETA
Sbjct: 2   ETEAAVTATVTAATMGIGTRKRDLKPYKGIRMRKWGKWVAEIREPNKR-SRIWLGSYATP 60

Query: 182 EDAALAYDRAAFRMRGSRALLNFP 205
           E AA AYD A F +RG  A LNFP
Sbjct: 61  EAAARAYDTAVFYLRGPSARLNFP 84


>G10 (Amino Acid Sequence) (gf=1)
         Length = 138

 Score = 74.7 bits (182), Expect = 4e-15
 Identities = 38/71 (53%), Positives = 48/71 (67%), Gaps = 5/71 (7%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           K Y+G+R R WGK+ AEIR+P K  +R+WLGSYKTAVAAARAYDTAVFYLRGPSARLNFP
Sbjct: 22  KRYKGIRMRKWGKWVAEIREPNKR-SRIWLGSYKTAVAAARAYDTAVFYLRGPSARLNFP 80

Query: 206 LRV----NSGE 212
             V    N GE
Sbjct: 81  EEVFKDGNGGE 91


>G1387 (Amino Acid Sequence) (gf=1)
         Length = 189

 Score = 73.9 bits (180), Expect = 6e-15
```

```
 Identities = 42/85 (49%), Positives = 52/85 (60%), Gaps = 2/85 (2%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           K +RGVRQR WG + +EIR P      RVWLGTF+TAE AA AYD+AA  M G  A  NFP
Sbjct: 5   KKFRGVRQRQWGSWVSEIRHPLLK-RRVWLGTFDTAETAARAYDQAAVLMNGQSAKTNFP 63

Query: 206 LRVNSGEPDDPVRIKSKRSSFSSSNE 230
           + + S   + + I S   S  S +E
Sbjct: 64  V-IKSNGSNSLEINSALRSPKSLSE 87


>G1386 (Amino Acid Sequence) (gf=1)
         Length = 194

 Score = 73.9 bits (180), Expect = 6e-15
 Identities = 37/77 (48%), Positives = 50/77 (64%), Gaps = 1/77 (1%)

Query: 129 KKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAY 188
           KK++          +K  ++RGVR R WGK+ +EIR+P K  +R+WLGTF TAE AA A+
Sbjct: 26  KKKRAKDDDDEKVVSKHPNFRGVRMRQWGKWVSEIREPKKK-SRIWLGTFSTAEMAARAH 84

Query: 189 DRAAFRMRGSRALLNFP 205
           D AA  ++G  A LNFP
Sbjct: 85  DVAALAIKGGSAHLNFP 101


>G1379 (Amino Acid Sequence) (gf=1)
         Length = 184

 Score = 73.2 bits (178), Expect = 1e-14
 Identities = 35/70 (50%), Positives = 47/70 (67%), Gaps = 1/70 (1%)

Query: 136 VSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRM 195
           V+   T  + + Y+G+R R WGK+ AEIR+P K  +R+WLG++ T E AA AYD A F +
Sbjct: 9   VAVPGTRKRDRPYKGIRMRKWGKWVAEIREPNKR-SRLWLGSYSTPEAAARAYDTAVFYL 67

Query: 196 RGSRALLNFP 205
           RG  A LNFP
Sbjct: 68  RGPTARLNFP 77


>G441 (Amino Acid Sequence) (gf=1)
         Length = 244

 Score = 73.2 bits (178), Expect = 1e-14
 Identities = 34/58 (58%), Positives = 44/58 (75%), Gaps = 1/58 (1%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           YRGVR R WGK+ +EIR+P K  +R+WLGTF T E AA A+D AA  ++G+ A+LNFP
Sbjct: 43  YRGVRMRTWGKWVSEIREPRKK-SRIWLGTFSTPEMAARAHDAAALTIKGTSAVLNFP 99


>G39 (Amino Acid Sequence) (gf=1)
         Length = 187

 Score = 73.2 bits (178), Expect = 1e-14
 Identities = 38/84 (45%), Positives = 54/84 (64%), Gaps = 5/84 (5%)

Query: 126 VPVKKEKTSPVSAAVTAA----KGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETA 181
           +P    K  +S+++  A     K K YRGVR+R WGK+ +EIR P K G R+WLG++ TA
Sbjct: 1   MPPSPPKSPPISSSLKGAHEDRKFKCYRGVRKRSWGKWVSEIRVP-KTGRRIWLGSYDAP 59

Query: 182 EDAALAYDRAAFRMRGSRALLNFP 205
           E AA AYD A F +RG + + NFP
Sbjct: 60  EKAARAYDAALFCIRGEKGVYNFP 83
```



```
>G2583 (Amino Acid Sequence) (gf=1)
         Length = 189

 Score = 72.8 bits (177), Expect = 1e-14
 Identities = 38/72 (52%), Positives = 46/72 (63%), Gaps = 1/72 (1%)

Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           + +RGVRQR WG + +EIR P      RVWLGTFETAE AA AYD+AA  M G  A  NFP
Sbjct: 5   RKFRGVRQRQWGSWVSEIRHPLLK-RRVWLGTFETAEAAARAYDQAALLMNGQNAKTNFP 63

Query: 206 LRVNSGEPDPVR 217
           +  +     D V+
Sbjct: 64  V-VKSEEGSDHVK 75


>G1846 (Amino Acid Sequence) (gf=1)
         Length = 221

 Score = 72.8 bits (177), Expect = 1e-14
 Identities = 37/74 (50%), Positives = 51/74 (68%), Gaps = 2/74 (2%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPPLR 207
           YRGVR R WGK+ +EIR+P K  +R+WLGTF TAE A  A+D AA  ++G+ A+LNFP
Sbjct: 19  YRGVRMRSWGKWVSEIREPRKK-SRIWLGTFPTAEMAMRAHDVAAMSIKGTSAILNFP-E 76

Query: 208 VNSGEPDPVRIKSK 221
           ++   P PV +  +
Sbjct: 77  LSKLLPRPVSLSPR 90


>G1755 (Amino Acid Sequence) (gf=1)
         Length = 174

 Score = 72.4 bits (176), Expect = 2e-14
 Identities = 44/103 (42%), Positives = 60/103 (57%), Gaps = 14/103 (13%)

Query: 115 ETPESPAAVD---SVPVKKEKTSPVSAAVTAAKGK--------HYRGVRQRPWGKFAAEI 163
           E  E+ + +D   S P++K    P  +    KGK        Y GVRQR WGK+ AEI
Sbjct: 32  EQTEADSCIDGGGSKPIRK--APPKRSRKGCMKGKGGPENGICDYTGVRQRTWGKWVAEI 89

Query: 164 RDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPL 206
           R+P + GA++WLGTF ++ +AA AYD A+  + G  A LN PL
Sbjct: 90  REPGR-GAKLWLGTFSSSYEAALAYDEASKAIYGQSARLNLPL 131


>G1277 (Amino Acid Sequence) (gf=1)
         Length = 153

 Score = 72.0 bits (175), Expect = 2e-14
 Identities = 34/67 (50%), Positives = 46/67 (67%), Gaps = 1/67 (1%)

Query: 139 AVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGS 198
           AV   + + ++G+R R WGK+ AEIR+P K  +R+WLG++ T E AA AYD A F +RG
Sbjct: 12  AVKMKRERPPKGIRMRKWGKWVAEIREPNKR-SRLWLGSYSTPEAAARAYDTAVFYLRGP 70

Query: 199 RALLNFP 205
            A LNFP
Sbjct: 71  TATLNFP 77


>G21 (Amino Acid Sequence) (gf=1)
         Length = 295
```

```
 Score = 72.0 bits (175), Expect = 2e-14
 Identities = 34/58 (58%), Positives = 44/58 (75%), Gaps = 1/58 (1%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           YRGVR R WGK+ +EIR+P K  +R+WLGT+ TAE AA A+D AA  ++G+ A LNFP
Sbjct: 100 YRGVRMRSWGKWVSEIREPRKK-SRIWLGTYPTAEMAARAHDVAALAIKGTTAYLNFP 156


>G2573 (Amino Acid Sequence) (gf=1)
         Length = 307

 Score = 71.6 bits (174), Expect = 3e-14
 Identities = 38/74 (51%), Positives = 49/74 (65%), Gaps = 7/74 (9%)

Query: 130 KEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYD 189
           K K  P +A  T       +RGVRQR WGK+ AEIR+P + G R+WLGTF T+ +AA+AYD
Sbjct: 22  KGKGGPENATCT------FRGVRQRTWGKWVAEIREPNR-GTRLWLGTFNTSVEAAMAYD 74

Query: 190 RAAFRMRGSRALLN 203
            AA ++ G  A LN
Sbjct: 75  EAAKKLYGHEAKLN 88


>G27 (Amino Acid Sequence) (gf=1)
        Length = 179

 Score = 71.6 bits (174), Expect = 3e-14
 Identities = 39/79 (49%), Positives = 49/79 (61%), Gaps = 5/79 (6%)

Query: 131 EKTSPVSAAVTAAKG----KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAAL 186
           EKT  S      KG     ++RGVR R WGK+ +EIR+P K  +R+WLGTF T E AA
Sbjct: 19  EKTGKSSKTKNEQKGVSKQPNFRGVRMRQWGKWVSEIREPRKK-SRIWLGTFSTPEMAAR 77

Query: 187 AYDRAAFRMRGSRALLNFP 205
           A+D AA  ++G  A LNFP
Sbjct: 78  AHDVAALAIKGGSAHLNFP 96


>G1381 (Amino Acid Sequence) (gf=1)
         Length = 256

 Score = 71.2 bits (173), Expect = 4e-14
 Identities = 37/74 (50%), Positives = 49/74 (66%), Gaps = 2/74 (2%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLR 207
           YRGVR R WGK+ +EIR+P K  +R+WLGT+ TAE AA A+D AA  ++G+    LNFP
Sbjct: 71  YRGVRMRSWGKWVSEIREPRKK-SRIWLGTYPTAEMAARAHDVAALAIKGNSGFLNFP-E 128

Query: 208 VNSGEPDPVRIKSK 221
           ++   P PV    K
Sbjct: 129 LSGLLPRPVSCSPK 142


>G2510 (Amino Acid Sequence) (gf=1)
         Length = 192

 Score = 70.9 bits (172), Expect = 5e-14
 Identities = 40/95 (42%), Positives = 57/95 (59%), Gaps = 13/95 (13%)

Query: 123 VDSVPVKKEKTSPVSAAVTAAKG----------KH--YRGVRQRPWGKFAAEIRDPAKNG 170
           + S PV   + +  ++AV +  G           +H  Y GVR+R WGK+ +EIR+P K
Sbjct: 8   LSSPPVTNNEPTATASAVKSCGGGGKETSSSTTRHPVYHGVRKRRWGKWVSEIREPRKK- 66

Query: 171 ARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           +R+WLG+F   E AA AYD AAF ++G +A LNFP
```



```
Sbjct: 67  SRIWLGSFPVPEMAAKAYDVAAPCLKGRKAQLNFP 101


>G2291 (Amino Acid Sequence) (gf=1)
         Length = 256

 Score = 70.5 bits (171), Expect = 7e-14
 Identities = 34/57 (59%), Positives = 44/57 (76%), Gaps = 1/57 (1%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNF 204
           + GVR+RPWG+++AEIRD      R WLGTF+TAE+AA AYD AA R+RG++A  NF
Sbjct: 116 FMGVRKRPWGRWSAEIRDRIGR-CRHWLGTFDTAEEAARAYDAAARRLRGTKAKTNF 171


>G2108 (Amino Acid Sequence) (gf=1)
         Length = 219

 Score = 69.7 bits (169), Expect = 1e-13
 Identities = 35/57 (61%), Positives = 42/57 (73%), Gaps = 1/57 (1%)

Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNF 204
           Y GVR+RPWG++AAEIR+P     R WLGTF+TAE+AA AYD AA  + GS A  NF
Sbjct: 21  YLGVRRRPWGRYAAEIRNPFTK-ERHWLGTFDTAEEAAFAYDVAARSISGSLATTNF 76


>G3 (Amino Acid Sequence) (gf=1)
         Length = 153

 Score = 69.7 bits (169), Expect = 1e-13
 Identities = 35/63 (55%), Positives = 44/63 (69%), Gaps = 1/63 (1%)

Query: 143 AKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALL 202
           A  K YRG+R+R WGK+ AEIR+P K  +R+WLG++ T   AA AYD A F +RG  A L
Sbjct: 26  ATRKPYRGIRRRKWGKWVAEIREPNKR-SRLWLGSYTTDIAAARAYDVAVFYLRGPSARL 84

Query: 203 NFP 205
           NFP
Sbjct: 85  NFP 87


>G13 (Amino Acid Sequence) (gf=1)
         Length = 253

 Score = 68.9 bits (167), Expect = 2e-13
 Identities = 38/77 (49%), Positives = 49/77 (63%), Gaps = 3/77 (3%)

Query: 130 KEKTSPVSAAVTAAKGKHYR--GVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALA 187
           + +T+P       +KG    +  GVRQRP GK+ AEI+D   +  R+WLGTFETAE+AA A
Sbjct: 2   EHQTTPKQKTKEKSKGNKTKFVGVRQRPSGKWVAEIKDTTQK-IRMWLGTFETAEEAARA 60

Query: 188 YDRAAFRMRGSRALLNF 204
           YD AA  +RGS    NF
Sbjct: 61  YDEAACLLRGSNTRTNF 77


>G2115 (Amino Acid Sequence) (gf=1)
         Length = 230

 Score = 68.2 bits (165), Expect = 3e-13
 Identities = 45/114 (39%), Positives = 57/114 (49%), Gaps = 25/114 (21%)

Query: 97  PSSSSSDEDRSSPPSVKIETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPW 156
           PSSSSS    SS  S K                K K S +          K Y+GVR R W
Sbjct: 23  PSSSSSSSSSSSSSCK--------------NKNKKSKI---------KKYKGVRMRSW 58
```

```
Query: 157 GKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNFPLRVNS 210
           G + +EIR P +   R+WLG++ TAE AA AYD A   ++G +A LNFP   +S
Sbjct: 59  GSWVSEIRAPNQK-TRIWLGSYSTAEAAARAYDVALLCLKGPQANLNFPTSSSS 111
```

>G869 (Amino Acid Sequence) (gf=1)
Length = 324

Score = 67.8 bits (164), Expect = 4e-13
Identities = 41/83 (49%), Positives = 53/83 (63%), Gaps = 5/83 (6%)

```
Query: 108 SFPSVKI-ETPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDP 166
           +FPS+++ E P   ++ DS   K +    VSA+    + K   GVRQR WGK+AAEIRDP
Sbjct: 75  NFPSMEVSEQPSESSSQDST--KTDGKIAVSASPAVPRKKPV-GVRQRKWGKWAAEIRDP 131

Query: 167 AKNGARVWLGTFETAEDAALAYD 189
            K   R WLGTF+T E+AA AYD
Sbjct: 132 IKK-TRTWLGTFDTLEEAAKAYD 153
```

>G40 (Amino Acid Sequence) (gf=1)
Length = 213

Score = 67.4 bits (163), Expect = 6e-13
Identities = 34/57 (59%), Positives = 42/57 (73%), Gaps = 1/57 (1%)

```
Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNF 204
           YRGVRQR  GK+ +E+R+P K   R+WLGTF+TAE AA A+D AA  +RG  A LNF
Sbjct: 48  YRGVRQRNSGKWVSEVREPNKK-TRIWLGTFQTAEMAARAHDVAALALRGRSACLNF 103
```

>G41 (Amino Acid Sequence) (gf=1)
Length = 207

Score = 66.6 bits (161), Expect = 1e-12
Identities = 34/57 (59%), Positives = 41/57 (71%), Gaps = 1/57 (1%)

```
Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFETAEDAALAYDRAAFRMRGSRALLNF 204
           YRGVRQR  GK+  E+R+P K   R+WLGTF+TAE AA A+D AA  +RG  A LNF
Sbjct: 42  YRGVRQRNSGKWVCELREPNKK-TRIWLGTFQTAEMAARAHDVAAIALRGRSACLNF 97
```

>G2067 (Amino Acid Sequence) (gf=1)
Length = 244

Score = 66.2 bits (160), Expect = 1e-12
Identities = 37/90 (41%), Positives = 48/90 (53%), Gaps = 1/90 (1%)

```
Query: 116 TPESFAAVDSVPVKKEKTSPVSAAVTAAKGKHYRGVRQRPWGKFAAEIRDPAKNGARVWL 175
           T  S +   S       TS +       K K Y+GVR R WG +  EIR P +   R+WL
Sbjct: 11  TTSSSSLSHSSSSSSSSTSALRHQSCKNKIKKYKGVRMRSWGSWVTEIRAPNQK-TRIWL 69

Query: 176 GTFETAEDAALAYDRAAFRMRGSRALLNFP 205
           G++ TAE AA AYD A   ++G +A LNFP
Sbjct: 70  GSYSTAEAAARAYDAALLCLKGPKANLNFP 99
```

Database: ptfgene
Posted date: Aug 27, 2002 10:23 AM
Number of letters in database: 551,722
Number of sequences in database: 1566

```
Lambda     K      H
   0.313    0.129    0.373
```



```
Gapped
Lambda     K      H
   0.267   0.0410    0.140
```

Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 325807
Number of Sequences: 1566
Number of extensions: 15185
Number of successful extensions: 306
Number of sequences better than 10.0: 151
Number of HSP's better than 10.0 without gapping: 148
Number of HSP's successfully gapped in prelim test: 3
Number of HSP's that attempted gapping in prelim test: 41
Number of HSP's gapped (non-prelim): 162
length of query: 268
length of database: 551,722
effective HSP length: 42
effective length of query: 226
effective length of database: 485,950
effective search space: 109824700
effective search space used: 109824700
T: 11
A: 40
X1: 16 (7.2 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 42 (21.9 bits)
S2: 49 (23.5 bits)